AD-A038 695 UTAH UNIV SALT LAKE CITY DEPT OF COMPUTER SCIENCE F/G 9/2

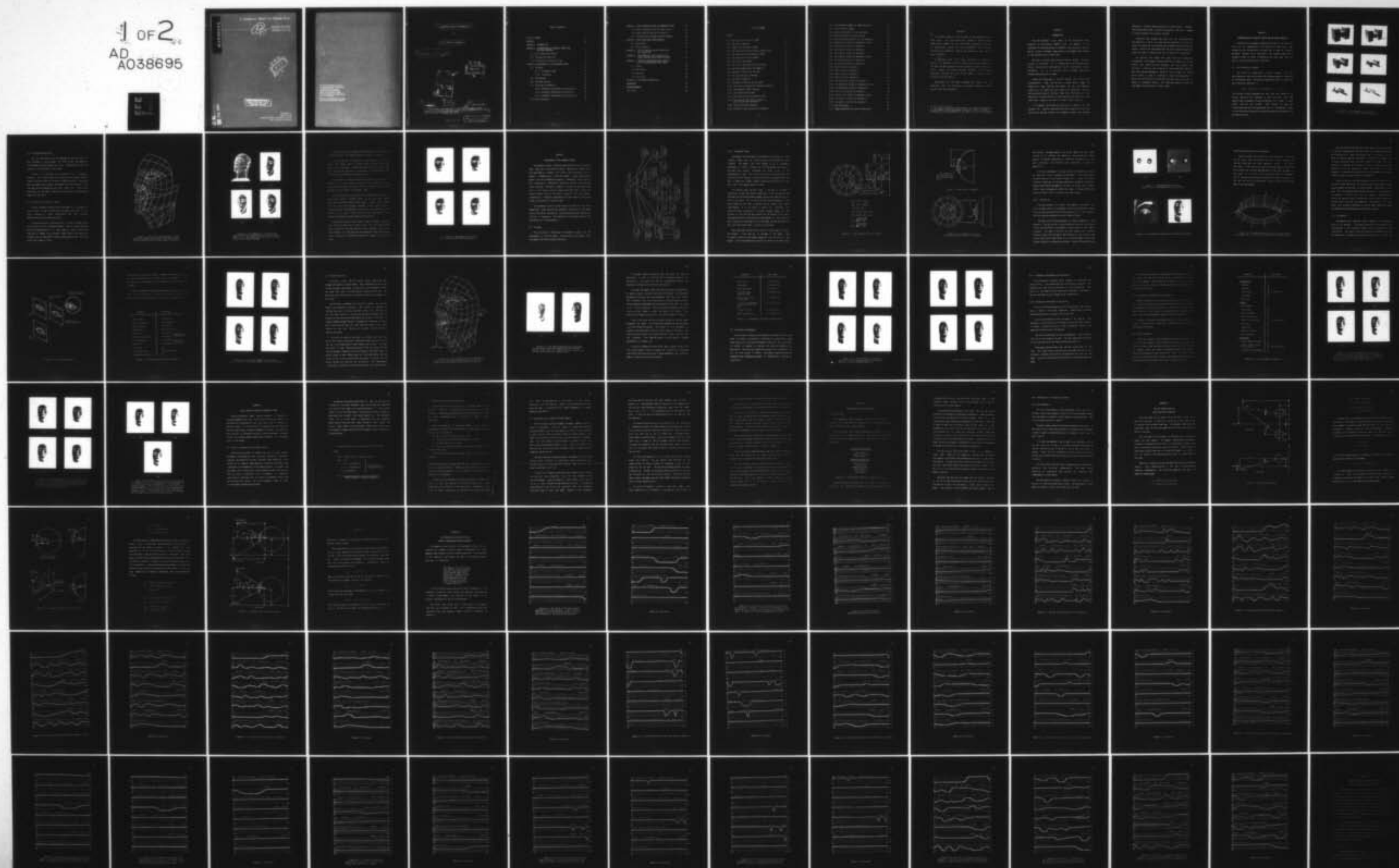
A PARAMETRIC MODEL FOR HUMAN FACES.(U)

DEC 74 F I PARKE DAHC15-73-C-0363

UNCLASSIFIED UTEC-CSC-75-047 NL

1 OF 2

AD A038695



A Parametric Model for Human Faces

12
NW

FREDERICK IRA PARKE

UNIVERSITY OF UTAH

DECEMBER 1974
UTEC-CSc-75-047
COMPUTER SCIENCE, UNIVERSITY OF UTAH
SALT LAKE CITY, UTAH 84112

6

A PARAMETRIC MODEL FOR HUMAN FACES

by

10 Frederick Ira Parke

9 *Technical rept.,*

11 Dec[redacted]74

12 108 p.

14 UTEC-CSc-75-047

15

This research was supported by the Advanced Research Projects Agency of the Department of Defense under Contract No. DAHC15-73-C-0363 and Contract No. F30602-70-C-0300.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

ABSTRACT *

This report presents a computer model for the representation of human faces. This three-dimensional, parametric model produces shaded facial images. The face, constructed of polygonal surfaces, is manipulated through the use of parameters which control interpolation, translation, rotation and scaling of the various facial features.

With this ~~is~~ model, very little input information is needed to specify and generate a specific face with a specific expression. The model has been successfully used to produce a large variety of facial images and several animated sequences. The animated sequences illustrate the power of the model to change facial expression and conformation.

Experience with the model indicates that fewer than 10 parameters must be manipulated to produce reasonable speech synchronized facial animation.

---

* This report reproduces a dissertation of the same title submitted to the Computer Science Department, University of Utah, in partial fulfullment of the requirements for the degree of Doctor of Philosophy.

CHAPTER 1

INTRODUCTION

The work presented in this report is the continuation and extention of my previous research [1,2]. In general, it is concerned with developing ways to represent, using polygons, objects having flexible surfaces. Specifically, this research has focused on the representation of human faces.

Why deal with human faces? There are several reasons. The face presents a challenge. It is flexible and varies from person to person. We have a well developed facility for interpreting facial expression. It is a difficult task to produce convincing computer-generated facial images.

Faces have always been a favorite subject for artists and photographers. They are skilled in using facial images to communicate ideas, emotions and moods. We can view computer graphics as an additional tool or medium for expression. As this medium matures we would certainly hope that techniques will be developed to handle objects with the complexity and subtlety of the human face. Hopefully this work is a step in that direction.

This research relies heavily on the work of Watkins [3] and Gouraud [4]. Watkins developed an efficent algorithm for solving the visible surface problem for polygonal objects and Gouraud

developed a "smooth" shading technique for these objects. Together these developments make it possible to generate "realistic" images of objects modeled with polygonal surfaces.

Some work has already been done with the two-dimensional representation of faces. Gillenson [5] has developed an interactive system for assisting "non-talented" users assemble and modify facial images. Chernoff [6] has demonstrated the use of computer-generated faces as a means of communicating "multi-dimensional" information.

My previous work showed that faces could be sucessfully represented with polygonal surfaces and that, at least to a limited extent, such faces could be animated. But the process was difficult, involving the collection of three-dimensional data for each face and each expression. We would like a model for faces which allows a class or universe of faces to be generated. A particular face from this universe would be specified by setting a number of parameter values. The remaining chapters discuss the development and application of such a model.

CHAPTER 2

INTERPOLATION AS A MEANS OF SPECIFYING POLYGONAL SURFACES

Interpolation of polygon surfaces was the basis of my earlier work with the representation and animation of faces [1,2]. The recognition that interpolation is a good way to specify flexible polygonal surfaces was a key factor in the research leading to a parametric model for faces. Interpolation has also been used to specify flexible surfaces other than faces.

2.1 The Interpolation Concept

The notion of interpolation is quite simple. In the one-dimensional case, we are given two values and asked to determine an intermediate value. The desired intermediate value is specified by a fractional coefficent α.

$$\text{value} = \alpha(\text{value 1}) + (1-\alpha)(\text{value 2}) \qquad 0 \leq \alpha \leq 1$$

This concept is easily expanded into more than one dimension by simply applying this procedure in each dimension. The idea generalizes to polygonal surfaces by applying the scheme to each vertex defining the surface. Each vertex will have two three-dimensional positions associated with it. Intermediate forms of the surface are achieved by interpolating each vertex between its extreme positions.

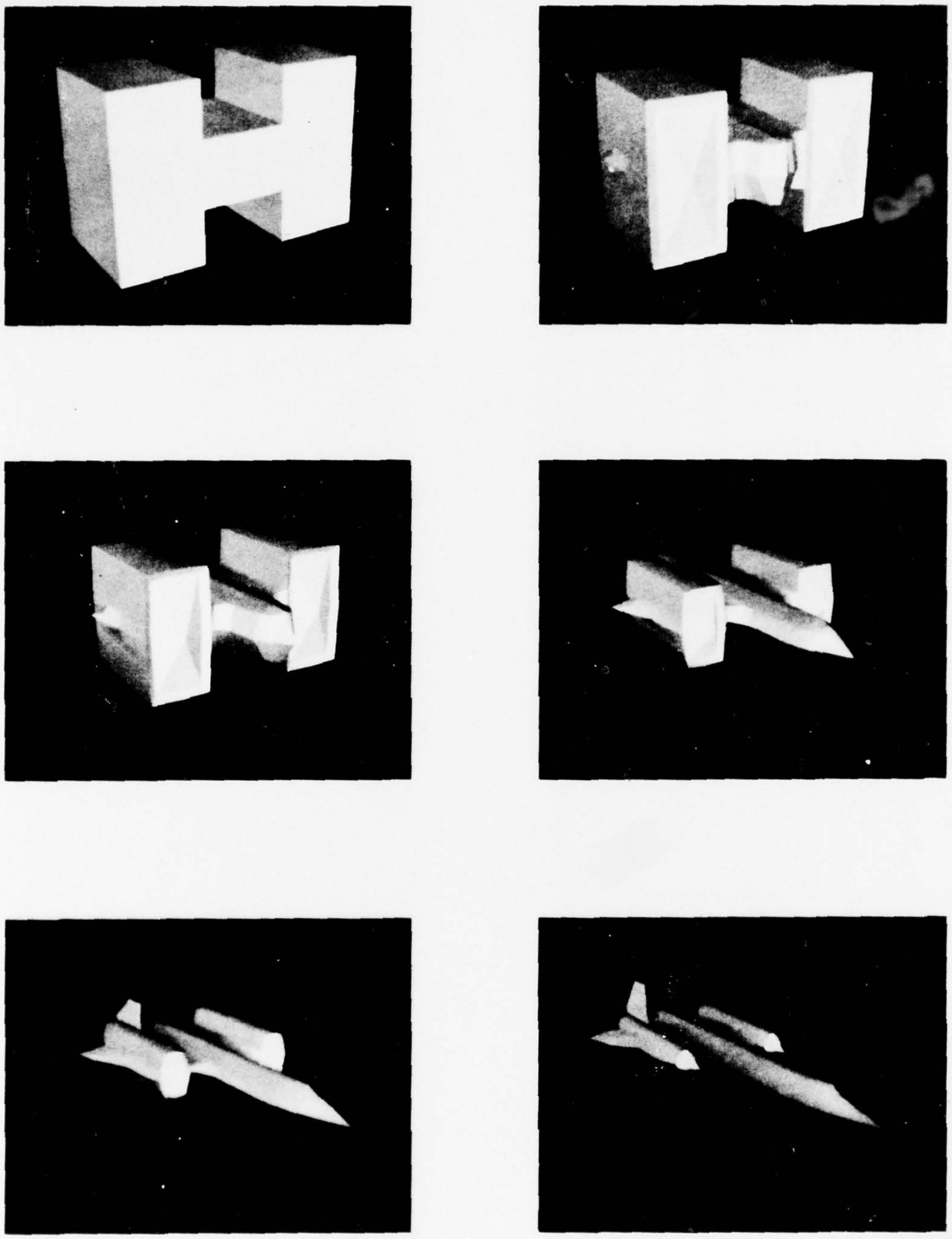


Figure 2.1 - The transition from a block
letter "H" to an SR71 aircraft using interpolation.

2.2 Topology Considerations

For this technique to work, the topology must be the same for both extremes of the surface. In other words, the number of vertices defining the surface and their interconnection must be identical for both forms of the surface.

Figure 2.1 illustrates the interpolation of a polygonal surface. This figure shows the transition of a surface from the shape of a block letter "H" to that of an SR71 aircraft. A topology was developed that allowed representation of the airplane. This topology was then mapped onto the block letter "H". Since both topologies are the same, interpolation can be used to transform one shape into the other.

2.3 Interpolation From Face to Face

A basic assumption underlying the development of a parametric facial model is that a single facial topology can be used. If the facial topology is fixed, manipulating the face involves manipulating only the vertex positions.

From previous work it was known that a fixed topology would allow a specific face to change expression. Would a single topology allow the representation of a wide range of faces? Could the topology be mapped onto different faces? Would the transition between faces be reasonable? To answer these questions we collected data from a number of faces.

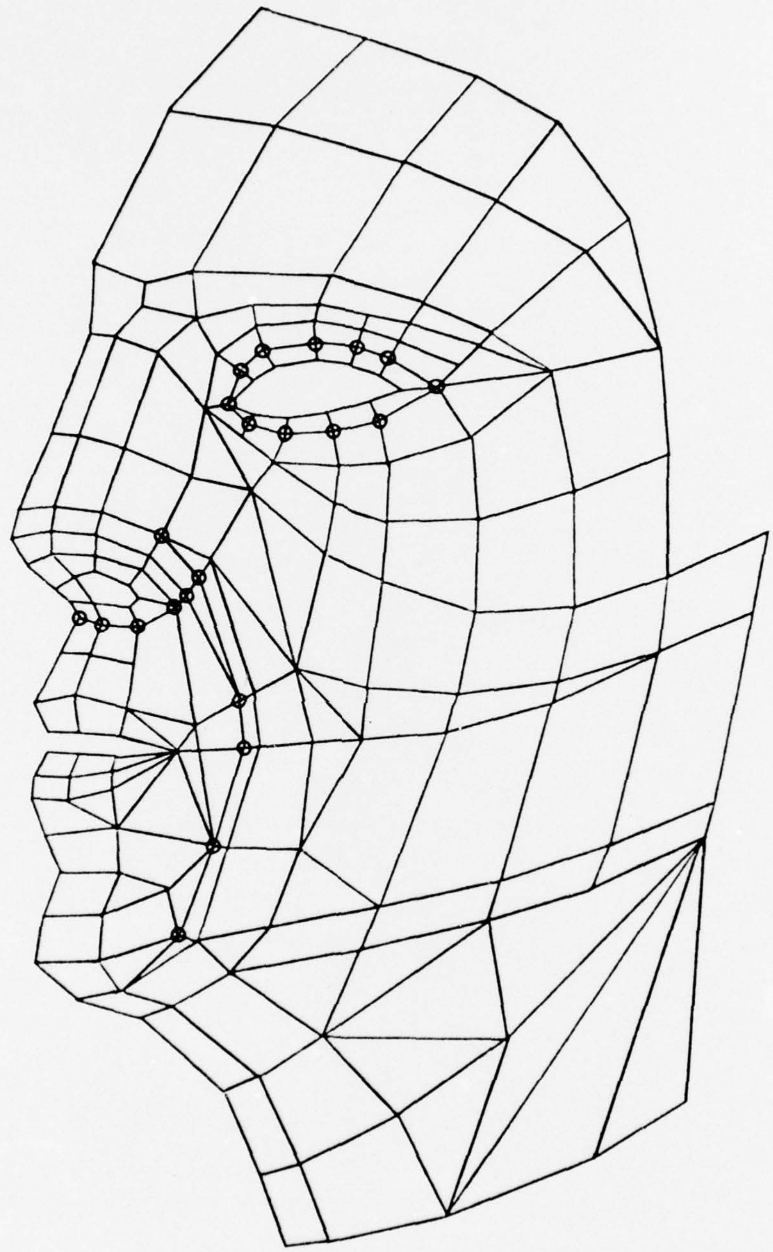
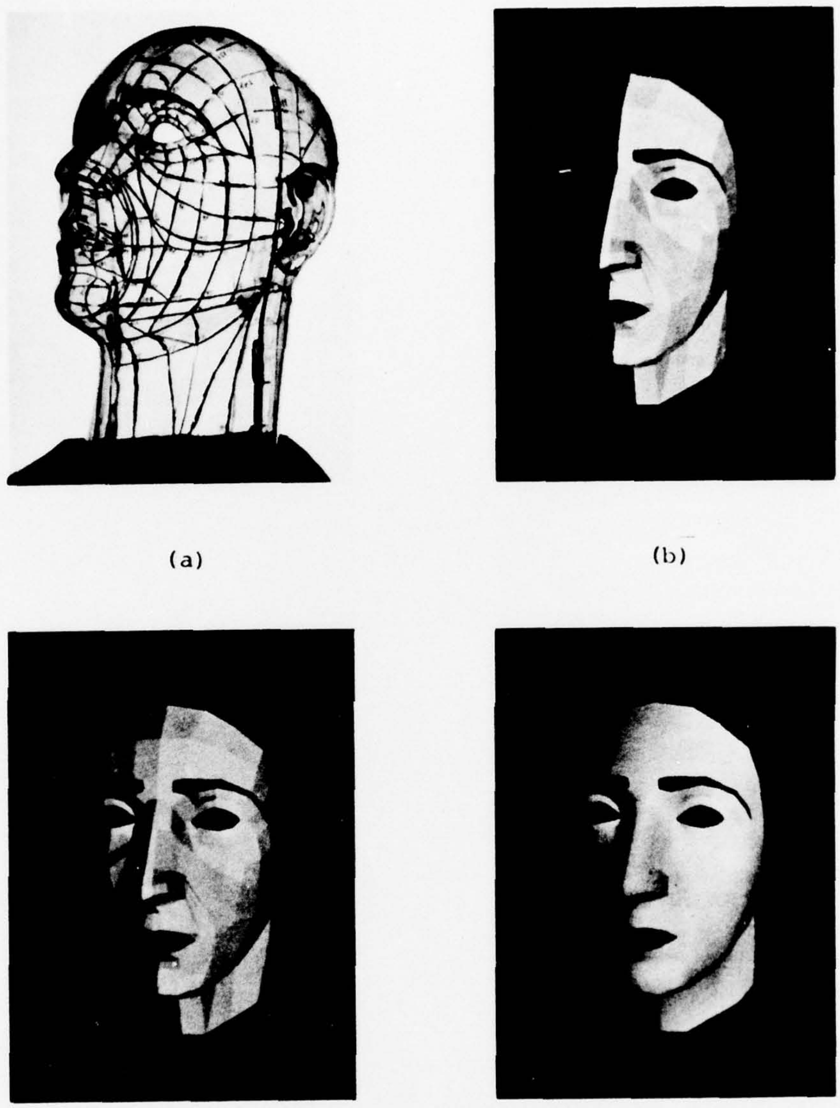


Figure 2.2 - The facial topology used to collect data from a number of real faces. The small circles indicate vertices that lie along creases in the face.



(a) (b)

(c) (d)

Figure 2.3 - The images (b), (c) and (d) were generated using data obtained from the plastic model shown in (a). The plastic model was used as a storage device for the facial topology.

The first step in this process was to determine a sufficiently flexible topology. The adopted topology is shown in Figure 2.2.

This topology was first applied to a plastic model of the human head. This model, shown in Figure 2.3(a), served as a storage device for the topology. The model was used as a guide each time the topology was applied to a real face. This assured that the topology would be identical from face to face.

The problem of collecting data from real faces is discussed in Appendix C. A more general discussion of this three-dimensional data collection process is given in [7]. A different but related approach to this problem is discussed in [8].

We collected data from one side of 10 different faces including the plastic model. Figures 2.3(b), (c) and (d) show images generated using data from the plastic model. Figure 2.3(b) shows half the face with polygonal shading. Figure 2.3(c) shows a complete faceted face generated by reflecting the data about the center of the face. Figure 2.3(d) shows the same symmetric face shaded using Gouraud's smooth shading technique [4].

Using data from the various faces, I made a computer animated film showing transitions from face to face. This film demonstrated that, at least for the faces used, a single topology would allow representation of many faces and the reasonable transition between faces. Figure 2.4 illustrates the transition between two of the faces.

Figure 2.4 - The transition from one face to another face using interpolation.

CHAPTER 3

DEVELOPMENT OF THE PARAMETRIC MODEL

The parametric model is based on data obtained from the plastic model described in the previous chapter. Manipulation capabilities were developed to transform this static data structure into a dynamic, parametrically controlled model. These manipulation capabilities are implemented by means of parameters which control the interpolation, translation, rotation or scaling of the various facial features. The model is symmetric. Except for the eyes, one side of the face is a mirror image of the other side. Figure 3.1 gives an overall view of the model structure. For this model, interpolation is applied independently to local regions of the face rather than globally to the whole face.

The parameters controlling the face are divided into two main categories, those controlling expression manipulation and those controlling facial conformation. Only the eye and mouth regions are involved in expression. The remainder of this chapter details the implementation of the parameters.

3.1 The Eyes

The first step in developing the parametric model was the development of realistic eyes. This was done in two phases: first the eyeball and then the eyelid mechanism.

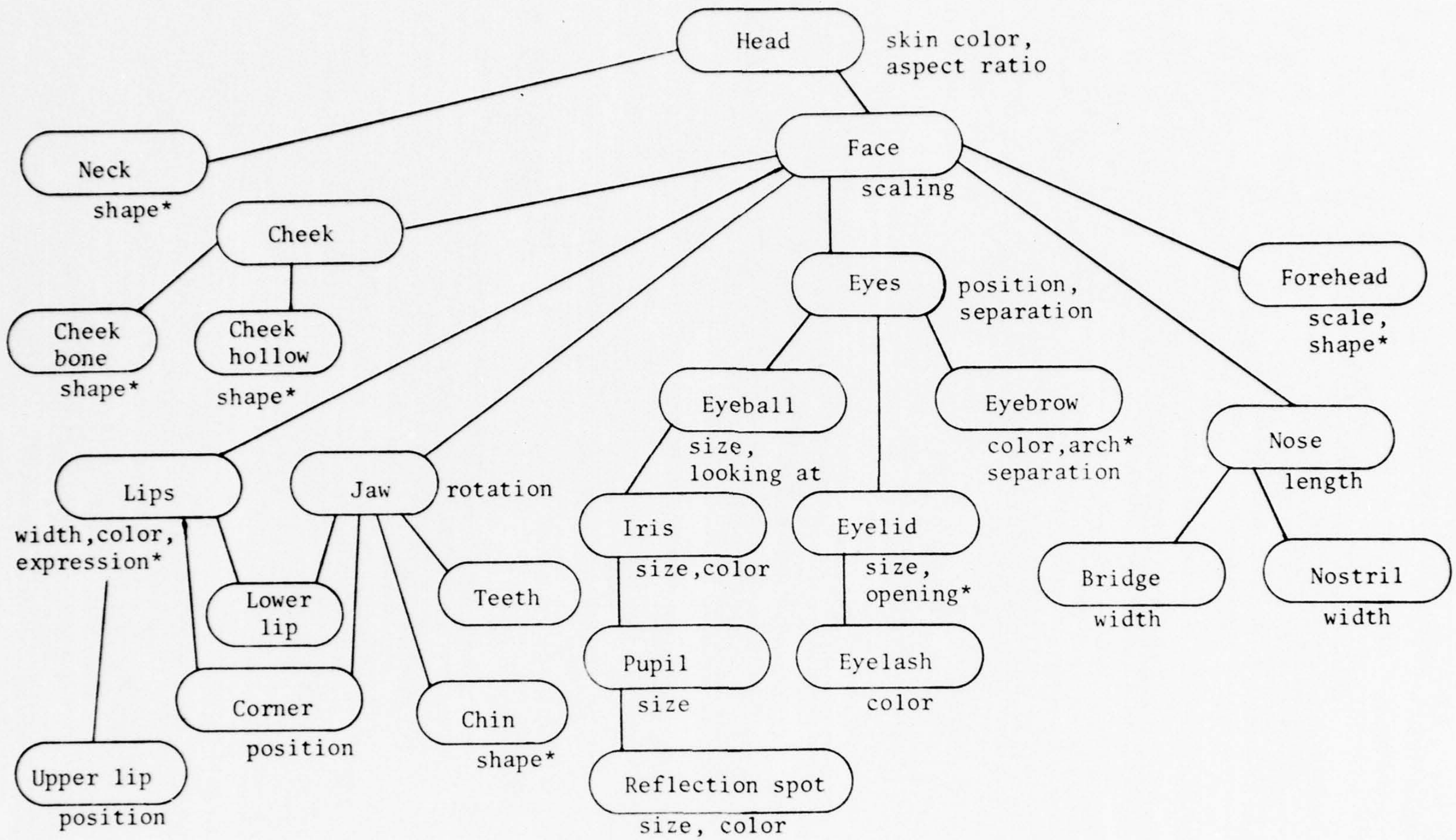


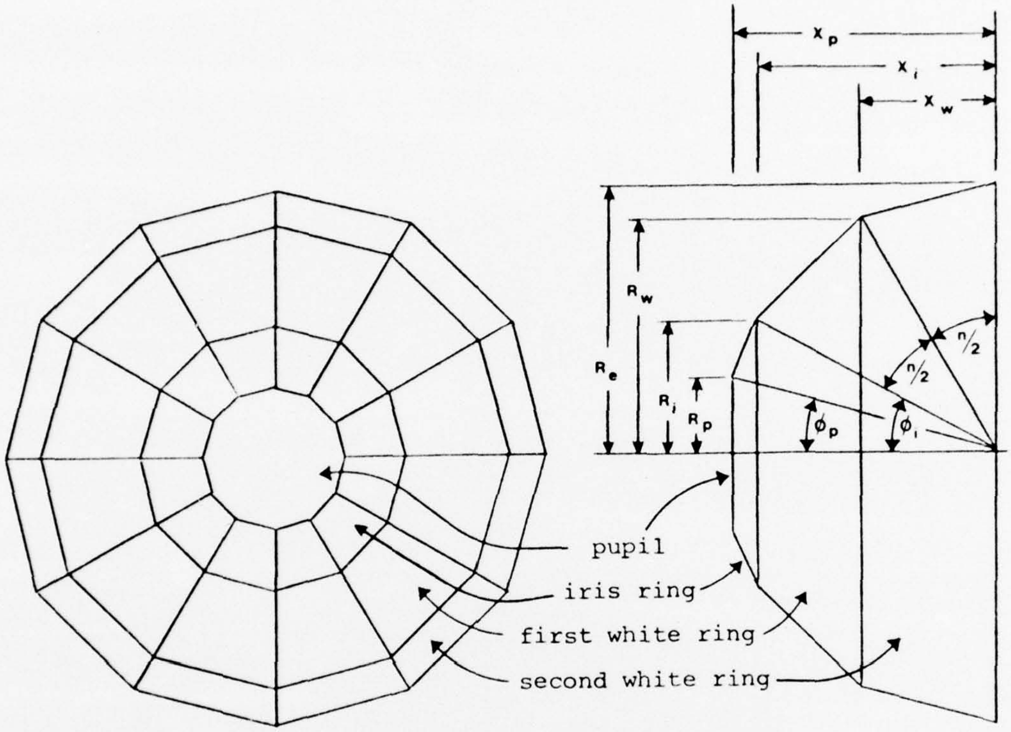
Figure 3.1 - The structure of the parametric model.
The parameters affecting the various nodes are shown. An *
indicates the use of interpolation to implement the parameter.

3.1.1 The Eyeball Model

The eyeball was developed as a procedure to be called for each instance rather than as a data structure to be scanned for each instance. The eyeball was first thought of as a polygonal hemisphere as shown in Figure 3.2. This hemisphere consists of a 12-sided pupil polygon surrounded by three rings of 12 quadralaterals each. The first ring is the iris and the other two form the white of the eye. The decision to use 12 polygons per ring was a compromise between eyeball complexity and a desire for the pupil and iris to appear nearly circular.

To achieve more realistic eyes, I decided to include a reflection of the light source on each eye. It was pointed out [9] that eye reflections are almost always visible over some portion of the iris or pupil. This is due to the fact that the eyeball is not really spherical but has a smaller partial sphere (the lens) superimposed on it as shown in Figure 3.3. Therefore, the reflection spot is modeled as a 6-sided polygon tangent to the surface of the lens and free to move over the surface of the lens. The exact position of the reflection spot depends on the position and orientation of the eyeball within the head, and the positions of the light source and viewer in relation to the head.

Real eyes have the ability to look at or track objects in their environment. This ability is included in the model. The orientation angles of each eyeball depend on the position of the eyeball within the head and the position of the point at which they



$$\varnothing_i = \sin^{-1} (R_i / R_e)$$

$$\varnothing_p = \sin^{-1} (R_p / R_e)$$

$$n = \pi/2 - \varnothing_i$$

$$R_w = R_e \sin (\varnothing_i + n/2)$$

$$X_w = R_e \cos (\varnothing_i + n/2)$$

$$X_i = \sqrt{R_e^2 + R_i^2}$$

$$X_p = \sqrt{R_e^2 + R_p^2}$$

Figure 3.2 - The construction of the eyeball.

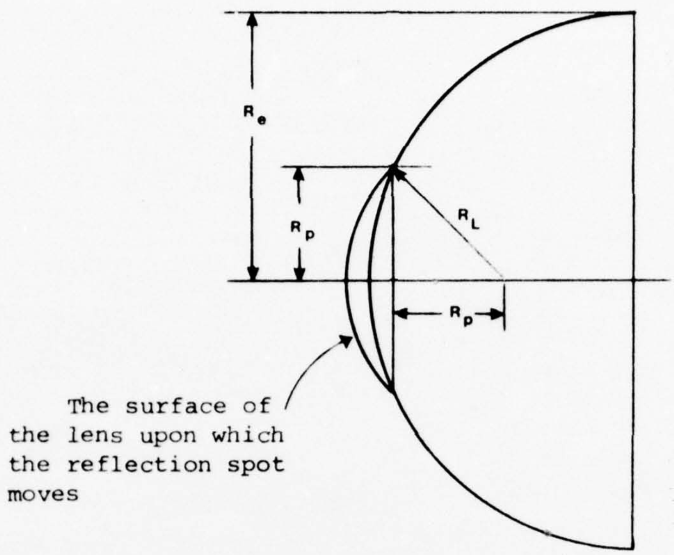


Figure 3.3 - The lens of the eyeball.

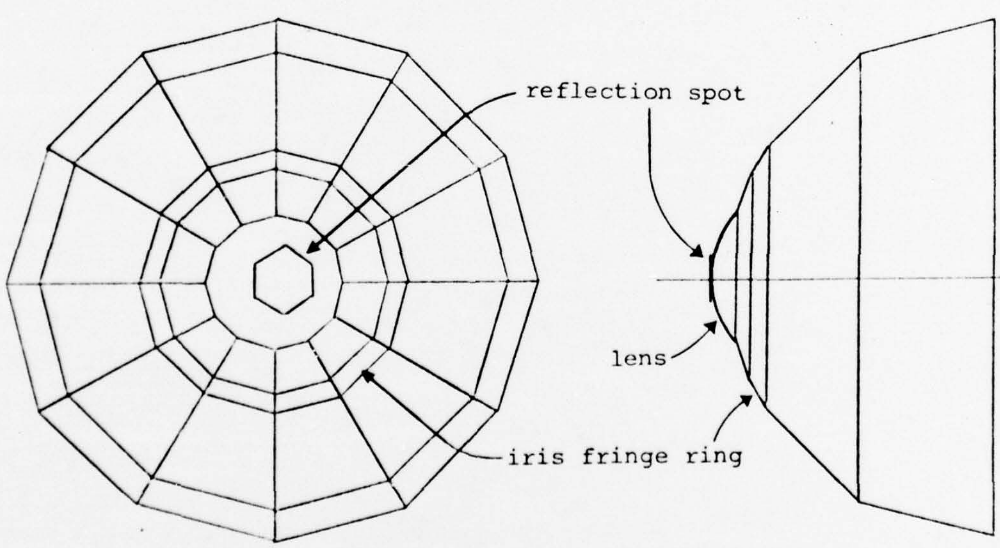


Figure 3.4 - The addition of an iris fringe and reflection spot to the eyeball.

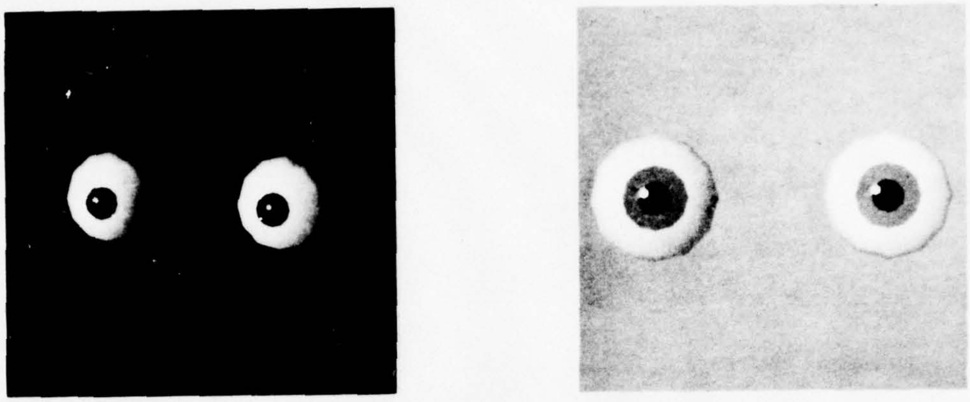
are looking. The eyes behave like normal eyes with both eyes looking at or tracking the same point. The orientation of each eyeball is computed independently. A detailed explanation of the eye orientation and reflection spot algorithms is given in Appendix A.

A further enhancement of the eye realism is obtained by adding an additional ring of polygons to the eyeball. This ring forms a fringe around the iris as shown in Figure 3.4. Figure 3.5 shows halftone renderings of the eyes produced by the eyeball procedure. Figure 3.5(a) shows the eyeballs with the iris fringe while Figure 3.5(b) shows the eyeballs without the fringe. In Figure 3.5(a) the relative sizes of the iris and pupil are nearly normal.

3.1.2 The Eyelids

The next problem is to "install" the eyeballs in the face. The first attempt was to fit the eyeballs into an existing static face. This was done by estimating the size and position of each eyeball and then generating eyeballs based on these estimates.

The major difficulty encountered is fitting the eyelids to the eyeballs. The eyelids are fitted by computing the polar coordinates (origin at the center of the eyeball) of each vertex of the eyelid polygons. The eyelid vertices are then mapped onto a sphere slightly larger than the eyeball and centered on it by setting the radius coordinate of each vertex to 1.1 times the eyeball radius and converting back to cartesian coordinates. Figure 3.6 shows halftone



(a) (b)

Figure 3.5 - Halftone renderings of the eyeballs with (a) and without (b) the iris fringe.

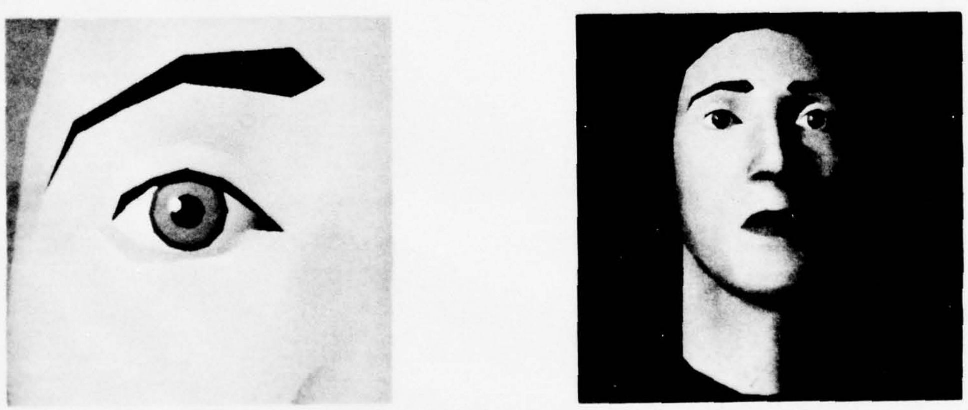


Figure 3.6 - The eyeballs "fitted" into the static facial model.

renderings obtained using this procedure.

There are several deficiencies with this technique. The most obvious being that the eyelids are still static. We want the eyelids to open and close under parametric control. The second deficiency is that the existing eyelid topology was not as detailed as we would like. We want some eyelashes, a few more polygons in the upper lid to allow it to follow the curvature of the eyeball better, especially when closed, and polygons to model the corner of the eye and the lip of the eyelids. Figure 3.7 shows the improved eyelid topology adopted.

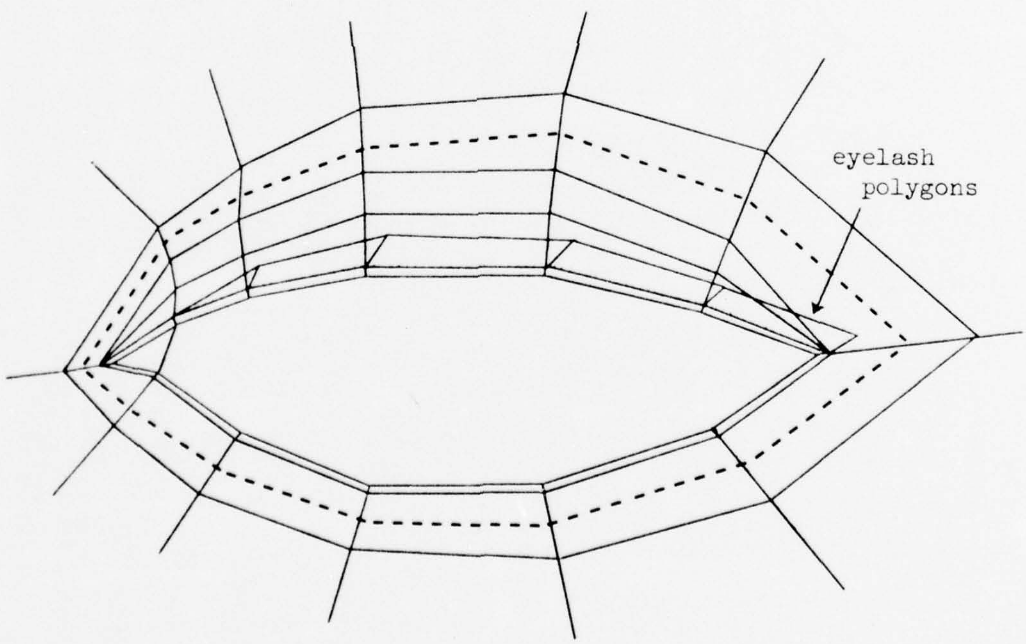


Figure 3.7 - The improved eyelid topology. Vertices within the dashed line are involved in opening and closing the eyelid.

With the improved topology we are now ready to face the problem of getting the eyelids to open and close. This problem was solved by a combination of linear interpolation and a variation of the spherical mapping idea described above. In Figure 3.7 a dashed line is shown enclosing a set of vertices. These vertices are the ones involved in opening and closing the eyelid. These vertices are defined in only two dimensions, height and width. The third dimension is obtained by projecting the vertices back onto a sphere slightly larger than and centered on the eyeball.

Using two sets of two-dimensional data values for these points, one set for open and one set for closed, we can interpolate for data values between the two sets. Projecting these interpolated values back onto the sphere produces the desired eyelid. This procedure, illustrated in Figure 3.8, gives the dynamic, parametrically controlled eyelid we need. The parameter controlling the eyelid is the one controlling the two-dimensional interpolation. In some sense this process is analogous to the real eyelid mechanism where two membranes are streched across a spherical surface.

3.2 The Eyebrows

The eyebrows are important facial features for expressing emotion and emphasis. The dynamic properties of the eyebrows are incorporated in the parametric model using interpolation and translation. Two sets of position values for the eyebrow vertices are specified. The eyebrows are varied from high to low arch by

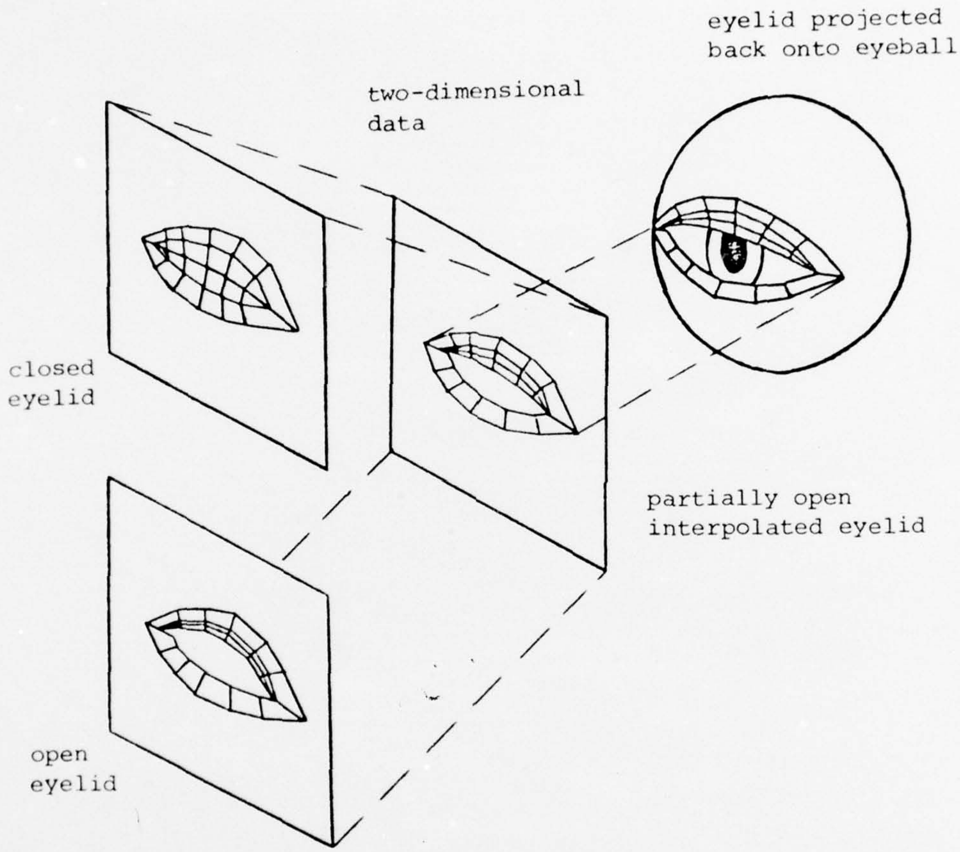


Figure 3.8 - The eyelid mechanism

interpolating between these values. An additional parameter allows horizontal translation of the inside end of the eyebrow. This parameter varies the separation of the eyebrows across the bridge of the nose.

A table of parameters affecting the eyes and eye region of the face is given in Figure 3.9. Figure 3.10 illustrates the effect of varying two of these parameters: eyelid opening and eyebrow arch.

Parameter	Value Range	
Eyebrow Arch	$0 \le value \le 1$	
Eyebrow Separation	$-25 \le value \le 25$	
Eyelid Opening	$0 \le value \le 1$	
Eyeball Size	$50 \le value \le 100$	
Iris Size	$0 \le value \le 1$	fraction of eyeball size
Reflection Spot Size	$0 \le value \le 1$	fraction of eyeball size
Iris Fringe Size	$0 \le value \le 1$	fraction of iris size
Pupil Size	$0 \le value \le 1$	fraction of iris size
Iris Color	any valid color	
Iris Fringe Color	any valid color	
Reflection Spot Color	any valid color	

Figure 3.9 - Parameters affecting the eye region.

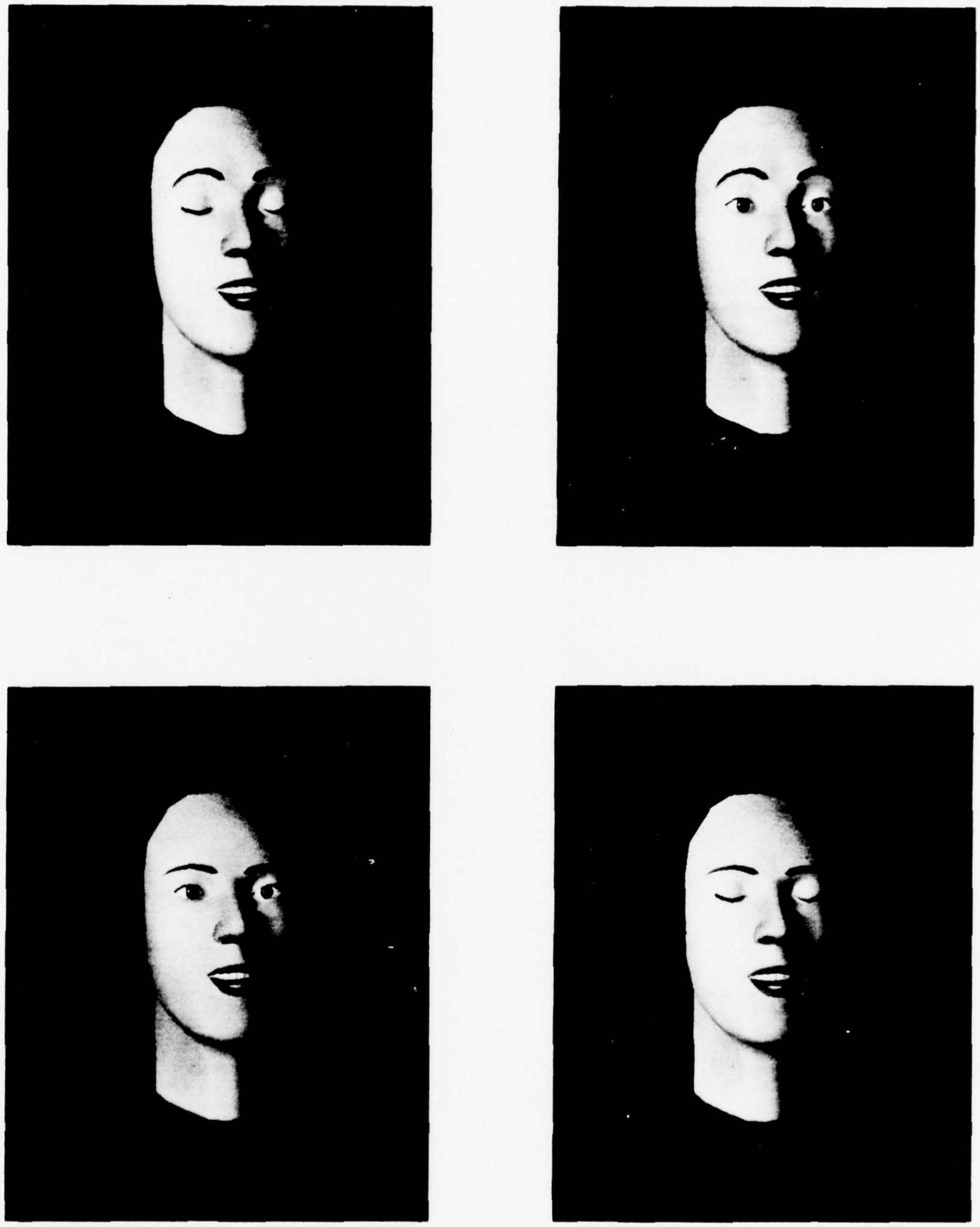


Figure 3.10 - Four facial images illustrating the
effect of two eye parameters: eyebrow arch and eyelid opening.

3.3 The Mouth and Jaw

To achieve a more realistic mouth, several additions and changes were made to the basis model. Teeth, modeled as a set of 32 4-sided polygons, were added. An additional row of polygons in both the upper and lower lip allows better curvature of the lips. An additional point at the corner of the mouth allows lip thickness at the corner.

Jaw rotation is necessary for the mouth to assume its various speech and expression positions. Jaw rotation is modeled by introducing an axis of rotation and then rotating the vertices of the jaw about this axis. The vertices between the dashed lines in Figure 3.11 are the vertices affected by jaw rotation. The axis of rotation passes through the point indicated and is parallel to the Y axis. Note that the lower lip, lower teeth and corner of the mouth rotate with the jaw. Positive jaw rotation has the effect of opening the mouth.

Initially, all points of the lower lip rotated with the jaw while the corner of the mouth rotated by an angle one-half that of the jaw. This scheme gives a rather square looking lower lip when the mouth is open. Rotating the center lower lip points with the jaw and gradually tapering the rotation angle of the other lower lip points gives a much better lower lip. Vertices farther from the center of the lip are rotated by smaller angles. The corner of the mouth rotates by one-third the jaw rotation. This improvement, illustrated in Figure 3.12, gives a more natural oval-looking mouth.

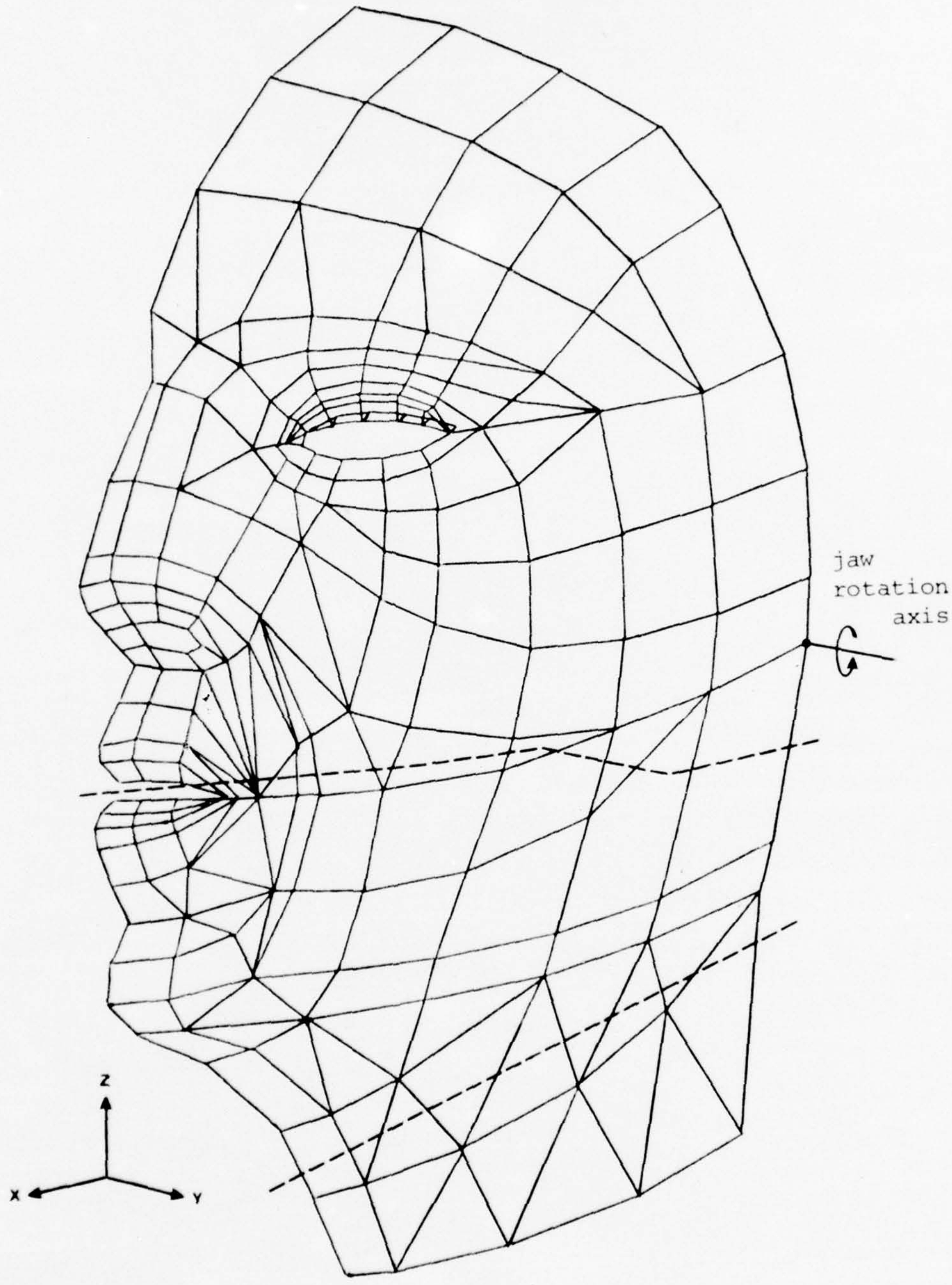


Figure 3.11 - The topology used for the parametric model. The jaw rotation axis is indicated. The dashed lines enclose the vertices effected by jaw rotation.

(a) (b)

Figure 3.12 - Two faces illustrating the improvement in mouth shape achieved by tapering the effect of jaw rotation on the lips. The original version is shown in (a) and the improved version is shown in (b).

In a dynamic model we certainly want the mouth to vary in expression. A step in this direction is to maintain data for two expressions. The mouth can then be interpolated between two expression extremes such as "smile" and "neutral."

The model has several additional mouth manipulation parameters. A scaling factor controls the width of the mouth. A translation parameter allows the lips to be moved away from the front teeth. The thickness of the lips at the corner of the mouth may be varied. Three translation parameters allow the corner of the mouth to move in all three dimensions. And a translation parameter allows the lower lip to be tucked up under the upper front teeth. This position is assumed by the mouth in forming the sounds "f" and "v".

Later, after experimenting with speech animation, another mouth parameter was added. This translation parameter allows the upper lip to be raised and lowered. The effect of this parameter is tapered from the center of the lip to the corner of the mouth. The center vertices receive full effect while the corner vertices are not affected. This tapering gives a more natural, rounded appearance to the upper lip.

A table of parameters affecting the mouth region along with their value ranges is shown in Figure 3.13. Figure 3.14 illustrates the effects obtained using three of these parameters: jaw rotation, upper lip position and mouth expression.

Parameter	Value Range
Jaw Rotation	$0 \leq$ value ≤ 20
Mouth Width	$.5 \leq$ value ≤ 1.5
Mouth Expression	$0 \leq$ value ≤ 1
Lip offset away from the teeth	$0 \leq$ value ≤ 30
Width of the corner of the mouth	$0 \leq$ value ≤ 25
X, Y and Z displacement of the corner of the mouth	$-25 \leq$ value ≤ 25
"f" and "v" tuck	$-20 \leq$ value ≤ 0
Upper lip position	$0 \leq$ value ≤ 20

Figure 3.13 - Parameters affecting the mouth region.

3.4 Conformation Parameters

Another group of manipulation parameters was added to allow the model to change in conformation. Conformation is used here to mean those features of the face that change or vary from one individual to another as opposed to features that vary from expression to expression. Some features change from expression to expression as well as from person to person. Here again, these conformation parameters are implemented by means of interpolation, scaling or translation.

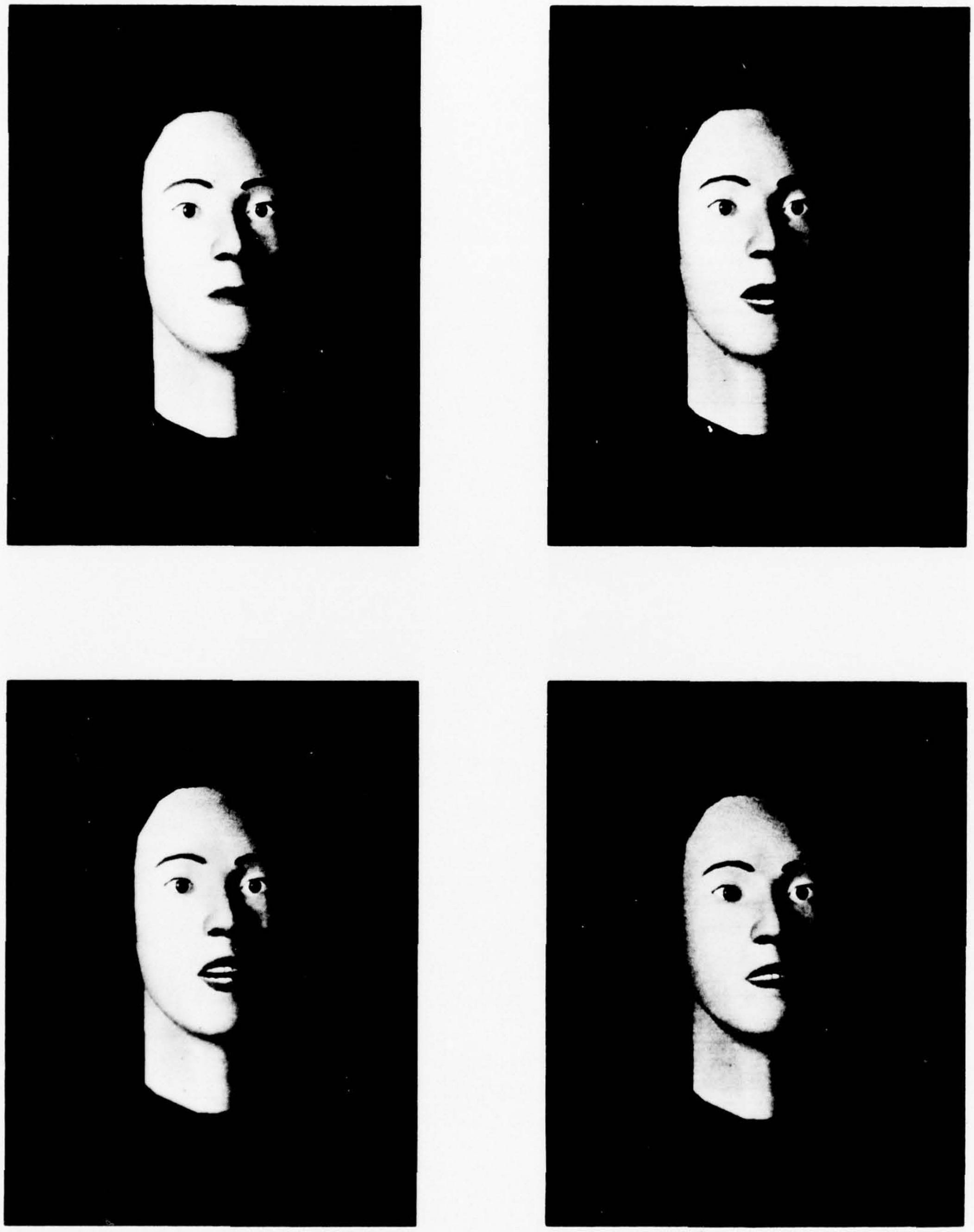


Figure 3.14 - A set of 8 faces illustrating the effect of three mouth parameters: jaw rotation, upper lip position and mouth expression.

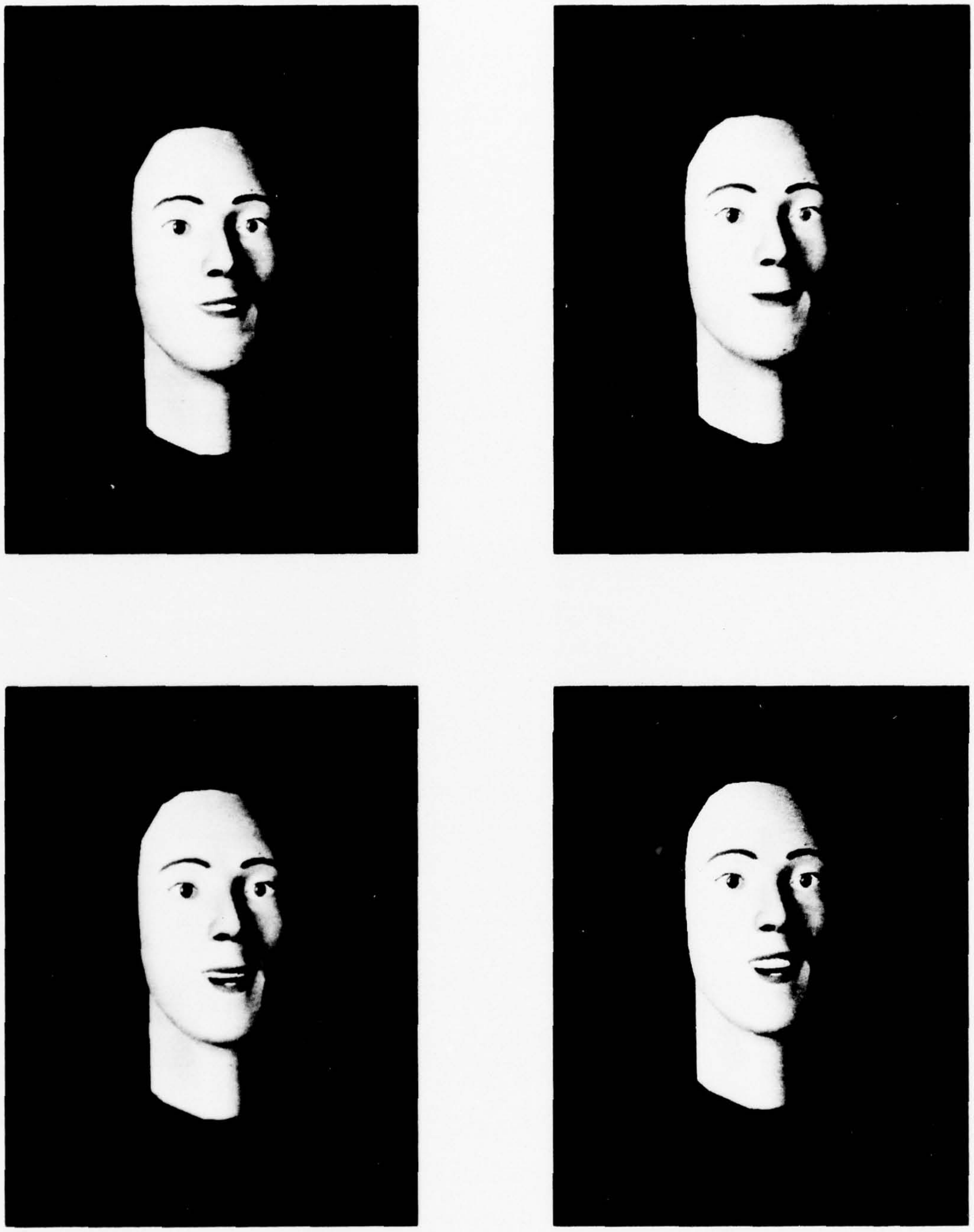


Figure 3.14 (continued)

3.4.1 Parameters Implemented by Interpolation

The conformation of several facial regions is controlled by interpolation. The forehead may vary from sloping to bulging. The cheekbone can range from not noticeable to very pronounced. The hollow of the cheek can vary from convex to concave. The shape of the chin and neck are also changed using interpolation.

3.4.2 Parameters Implemented using Scaling

A set of scaling parameters controls the scaling of the entire face in each of the three dimensions. Combinations of these parameters determine the aspect ratio of the face.

Two scaling parameters affect the shape of the eyelids. One parameter controls the width of the eyelid while the other controls its height. Varying the values of these parameters controls the shape and relative size of the eyelids.

The nose is affected by two scaling parameters. One controls the width of the bridge of the nose. The other determines the width of the lower portion of the nose including the nostrils.

Three scale factors control the vertical proportions of the face. One value controls the scaling of the area from the chin to the mouth. Another value controls the area from the chin to the eyes. The third value controls the scaling of the region above the eyes.

In addition there are scaling parameters affecting the width of the cheek area and the width of the jaw. The effect of the jaw scaling is tapered. The maximum effect is applied to the forward portion or point of the jaw. The scaling effect diminishes to zero at the rear of the jaw.

3.4.3 Parameters Implemented using Translation

Several parameters allow regions of the face to be displaced. It is possible to move the chin both up and down as well as forward and backward in relation to the rest of the face. Likewise the lower portion of the nose can be moved forward, backward, up or down. And finally the eyebrows may be displaced up or down.

A table of conformation parameters and their value ranges is given in Figure 3.15. Figure 3.16 illustrates the effect of some conformation parameters.

3.5 Other Parameters

There are a number of other parameters which affect either the face or its environment. These include the position of the eyeballs within the head, where the eyes are looking, where the light source is located, where the viewer is located and where he is looking. Additional parameters control the field of view, the color (grey level) of the various facial regions, the background grey level and the cosine power and diffuse factor used in shading the face.

Parameter	Value Range
Interpolation	
Forehead	$0 \le$ value ≤ 1
Cheekbone	
Cheek hollow	
Chin shape	
Neck shape	
Scaling	
Chin-to-mouth	$.5 \le$ value ≤ 1.5
Chin-to-eye	
Eye-to-forehead	
Eyelid X and Z size	
Head X, Y and Z scale	
Jaw width	
Cheek width	
Bridge of the nose	
End of the nose	
Translation	
X and Z offset for chin	$-50 \le$ value ≤ 50
X and Z offset for the end of the nose	
Z offset for the eyebrows	

Figure 3.15 - The conformation parameters.

(a)

(b)

(c)

(d)

Figure 3.16 - Effects achieved using some of the conformation parameters. The initial face is shown in (a). In (b) the forehead has been interpolated to a different shape. For (c) the neck was interpolated to a new shape. The jaw was scaled by a factor of .8 for (d).

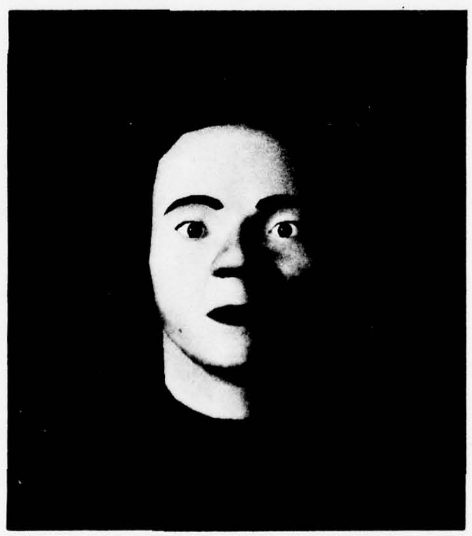
(e) (f)

(g) (h)

Figure 3.16(continued) - For (e) the chin-to-mouth scale was set to .85. For (f) the chin-to-mouth scale was .9 and the chin-to-eye scale was .8. In (g) the vertical scale for the head was increased to 1.15. In (h) the scale of the end of the nose was changed from 1.0 to .8.

(i)

(j)



(k)

Figure 3.16(continued) - In (i) the end of the nose was moved down. For (j) the horizontal scale of the head was set to .95 and the cheekbone interpolated to a slightly more prominent shape. Finally for (k) the horizontal scale of the head was changed to 1.1. The end of the nose was raised and scaled by 1.1. The bridge of the nose was narrowed and the eyebrows lowered. The cheekbones are slightly more prominent and the cheek hollows are a little more concave.

CHAPTER 4

FACIAL ANIMATION USING THE PARAMETRIC MODEL

Using the parametric model, facial animation is reduced to varying parameters over time. The difficult part of this task is to determine how the parameters must vary over time to achieve a desired effect. It is also difficult to specify and coordinate the usually large number of parallel parameter functions. The parallel specification of parameter functions is described below. A subset of the first problem, speech synchronized animation, is discussed later in the chapter.

4.1 Parallel Parameter Function Specification

When using this model for animation we want to vary several parameters simultaneously over any given time period. Since most computers are sequential, we must introduce some mechanism which allows us to specify these parallel functions and have them translated into a sequential list of specifications. A program was developed which has as input a set of parameter specifications and has as output a sequential list of commands. A parameter specification describes how the value of a parameter should vary over a given time interval. The list of commands is used as input to the program implementing the model.

The sequencing program accomplishes its task by storing the information from each parameter specification and then searching this data for each frame of the animated sequence. If the current frame is within the frame range of a specification then a command is generated to be included in the output command list. The parameter value for a given frame is determined from the parameter specification using the frame range, parameter value range, the current frame number and the parameter change function specified. The parameter change functions and the parameter value computation are given below.

$$\text{parameter value} = \text{initial value} + \text{PCF} * \text{DIFF}$$

where

$$\text{DIFF} = \text{final value} - \text{initial value}$$

$$\left.\begin{aligned} \text{PCF} &= \text{FF} \\ &\text{or } 1 - \text{COS}(\text{FF}*\pi)/2 \\ &\text{or } 1 - \text{COS}(\text{FF}*\pi/2) \\ &\text{or } \text{SIN}(\text{FF}*\pi/2) \end{aligned}\right\} \text{The parameter change functions}$$

$$\text{FF} = \frac{\text{current frame \# - initial frame \#}}{\text{final frame \# - initial frame \#}}$$

4.2 Speech Synchronized Facial Animation

Speech animation implies the ability to convey emphasis and emotion in addition to manipulating the mouth and lips. Speech animation is achieved by manipulating the facial parameters so that the facial expression and lip motions will match a spoken soundtrack.

What is required for lip animation? Madsen in his book on conventional animation [10] indicates the following capabilities are required for lip animation.

a) Open lips for the open vowels a, e and i

b) closed lips for the accent consonants p, b and m

c) an oval mouth for u, o and w

d) the ability to tuck the lower lip up under the upper front teeth for the f and v

e) and the ability to move between these lip positions as required.

The remaining sounds are formed mainly by the tongue and do not require precise animation. Madsen also indicates that realistic characters require more care in lip animation than abstract or non-realistic characters.

A number of the parameters detailed in the previous chapter are used to give the required lip animation capability. These parameters are jaw rotation, width of the mouth, displacement away from the teeth, translation of the lower lip up under the upper

front teeth and manipulation of the corner of the mouth. Manipulation of the eyebrows, eyelids and the expression of the mouth are used in conjunction with these parameters to convey emphassis and emotion.

4.3 The Production of Speech Animated Segments

The first step in producing speech animated segments was to obtain a soundtrack. Since we wished to compare the computer produced animation with the real speech, we filmed an actor reciting a poem. He was filmed using what is known as "double system". His image was recorded on film while his voice was recorded on magnetic tape. The end product of this procedure is a piece of movie film and a piece of sprocketed magnetic film. These two pieces are synchronized so that the sound corresponds frame for frame with the images on the movie film.

The next step was to analyze (search) the magnetic film for the various parts of speech i.e. open vowels, accent consonants, etc. The end product of this process is a frame by frame chart of the speech contained on the film.

A total of six computer generated speech segments were produced for use with this soundtrack. The first was produced in the following manner. Using the speech vs frame number chart and a mirror, I went through the poem mouthing the words. I watched my mouth motions in the mirror and translated them into parameter functions used to drive the model. Graphs of the parameter

functions used for this and the other segments are included in Appendix B. The parameters used for the mouth in this segment were jaw rotation, mouth expression, translation away from the teeth, mouth width, the f, v tuck and manipulation of the corner of the mouth. In addition there was some manipulation of the eyelids and the eyebrows.

The sequence generated was quite disappointing. For one thing, a fundamental problem with speech animation was encountered. Often the lip motions vary over very short time periods. They can move at about the same rate as the film frame rate. This means that a severe sampling problem occurs. How do you represent a motion that lasts only a frame or two? This segment was also over animated. This gave the impression that the lips were moving about twice as fast as the speech even though the accent consonants were properly synchronized.

For the second segment all lip or mouth motion was deleted except jaw rotation. The jaw rotation was "smoothed" out. By comparing the jaw rotation graphs for sequences 1 and 2, this smoothing can be seen. The resulting animated segment was again disappointing. The goal of simplifying and slowing the lip motions was achieved but the motion was not natural. The motion did not seem to match the speech very well even though the accent consonants were in proper synchronization.

For the third sequence, "f" and "v" tucks were added. This made essentially no difference in the results. The "f" and "v"

tucks occur over such short intervals that they are not visible.

For the fourth try, the width of the mouth was also varied. This gave some improvement, but the animation was still not right.

At this point, confidence in the model was somewhat diminished. Was the poor animation due to the parameter functions driving the model or was there a basic deficiency in the model itself? To answer this question, I decide to analyze, frame by frame, the film of the actor speaking. The results of this analysis would then be used to determine the parameter functions used to drive the model.

During the course of the analysis it became apparent that the model lacked an important capability. When the actor spoke he used his upper lip. The model, however, did not allow the upper lip to move. The model was modified to include this capability.

A fifth try at the speech animation was made using the improved model and the results of the frame by frame analysis. In this sequence the lip animation was much more convincing.

For the sixth try, the film was again analyzed frame by frame. This second, independent analysis was used to drive the model for this final sequence. Much more attention was given to the eyes, eyebrows and mouth expression in an effort to get more emotion into the animation. This final segment is quite convincing. Most viewers agree that it is at least on a par with with most conventional speech animation.

CHAPTER 5

CONCLUSIONS AND FUTURE RESEARCH

5.1 Conclusions

The parametric model developed during this research is certainly not "the" model for faces but it is a viable model and a good starting point for future research.

With this model, fewer than 10 parameters are needed to do a reasonable job of speech animation. The parameters found most effective for facial expression and lip animation are listed in Figure 5.1.

<u>LIP ANIMATION PARAMETERS</u>

Jaw rotation
Upper lip position
Mouth width

<u>EXPRESSION PARAMETERS</u>

Mouth expression
Eyebrow arch
Eyebrow separation
Eyelid opening
Pupil size
Eye tracking

Figure 5.1 - The parameters needed for speech animation.

Speech animation for a model with this level of realism is difficult. The faces are real enough that the viewer expects them

to behave realistically. The simplified techniques used in most character speech animation are not good enough. The "f" and "v" tucks are particularly useless.

The conformation parameters of the model are not as easily evaluated as the speech and expression parameters. These parameters do allow the conformation of the face to change. But, it is not clear exactly what conformation parameters are desirable. There are a number of additional conformation controls that might be tried. Parameters that might be tried include shape of the nose, nostril size and position, lip shape and thickness, eyebrow shape and thickness, relative size and position of the face with relation to the head, and the size and shape of the teeth. Experience with a model having a wide range of conformation control would be necessary to determine the best set of conformation parameters.

One deficiency of the current model is that it is symmetric. Since real faces are not symmetric, the ability to handle non-symmetric conformation and expressions would be a significant improvement. Another deficiency is that the face does not pivot on the neck. When people speak, they tend to pivot their heads. The ability to pivot the head would greatly improve the realism of the speech animation.

Obviously the realism of the images would be greatly improved if the entire head were modeled rather than just the face and neck. The addition of ears, hair and maybe a tongue would enhance the model. The addition of hair presents a problem, however. How is

hair represented with polygonal surfaces?

5.2 Future Research

Efficient techniques are being developed which use surface patches rather than polygons as a basis for creating shaded images. The development of parametric models based on surface patches seems to be a promising area of research.

Automatic speech synchronized animation seems within reach. A marriage of the current speech recognition technology and the speech animation capability demonstrated by this research could lead to major results.

In a sense the parametric facial model is an instrument we do not yet know how to play. It initiates a new area that might be labeled "computer acting." A new set of skills and intuitions is needed. These would be analogous to the skills and intuitions a conventional animator develops allowing him to work effectively with his media.

The facial model might be useful to psychologists interested in expression and non-verbal communication. The model allows separation of expressions and actions not normally seen independently in real people.

The development of parametric graphical models for classes of objects is a very open-ended area of study. The development of this model for faces is just an intial step into this area.

APPENDIX A

THE EYE TRACKING AND EYE REFLECTION SPOT ALGORITHMS

The facial model has the capability for the eyes to look at a specified position. In addition, a reflection of the light source is visible on the surface of each eye. This appendix describes the algorithms used to orient the eyeballs and to place the reflection spots on the eyes.

The first step in this process is to determine the orientation angles for each eyeball. The eyeball generated by the eyeball procedure is centered at the origin of its coordinate system with the X axis passing through the center of the pupil. The orientation angles desired are those that will rotate the eyeball so that it will be looking in the desired direction when it is positioned in the face.

Refering to Figure A.1 we see that each eye has two orientation angles, α and β, associated with it. The and for each eye are computed independently. The following equations are used to determine these angles.

$$\alpha_R = \arctan\left((Y_T - Y_R)/(X_T - X_R)\right)$$

$$\alpha_L = \arctan\left((Y_T - Y_L)/(X_T - X_L)\right)$$

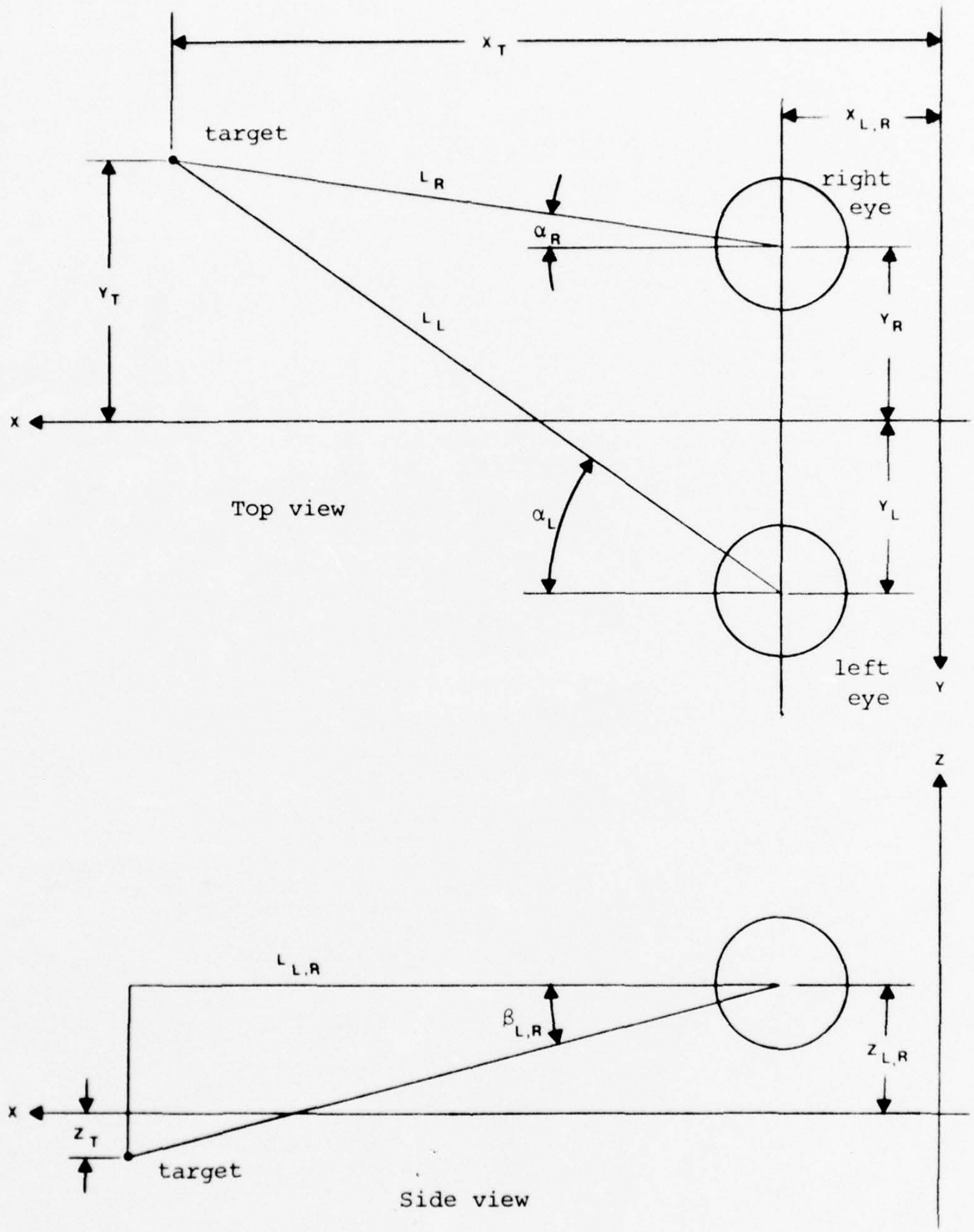


Figure A.1 - Eye orientation angles

$$\beta_R = \arctan((Z_R - Z_T)/L_R)$$

$$\beta_L = \arctan((Z_L - Z_T)/L_L)$$

$$L_R = ((Y_T - Y_R)^2 + (X_T - X_R)^2)^{1/2}$$

$$L_L = ((Y_T - Y_L)^2 + (X_T - X_L)^2)^{1/2}$$

The next step is to determine the position of the center of the lens sphere in the eyeball coordinate system. Refering to the upper left diagram in Figure A.2 we see that the center of the lens is displaced a distance L_c along the X axis of the eyeball. Assuming an eyeball of unit radius with the iris radius R_i specified as a fraction of this radius then the distance L_c is computed as follows.

$$1 = R_i^2 + (R_i + L_c)^2$$

$$1 - R_i^2 = (R_i + L_c)^2$$

$$L_c = (1 - R_i^2)^{1/2} - R_i$$

To find the true magnitude of L_c we need to multiply the equation above by the radius R_e of the eyeball.

$$L_c = ((1 - R_i^2)^{1/2} - R_i)\ R_e$$

Now what happens to the center of the lens as the eyeball is rotated by α and β into its proper orientation? Again refering to Figure A.2 we see that the center of the lens is now located at some position X_L, Y_L, Z_L in the eyeball coordinate system.

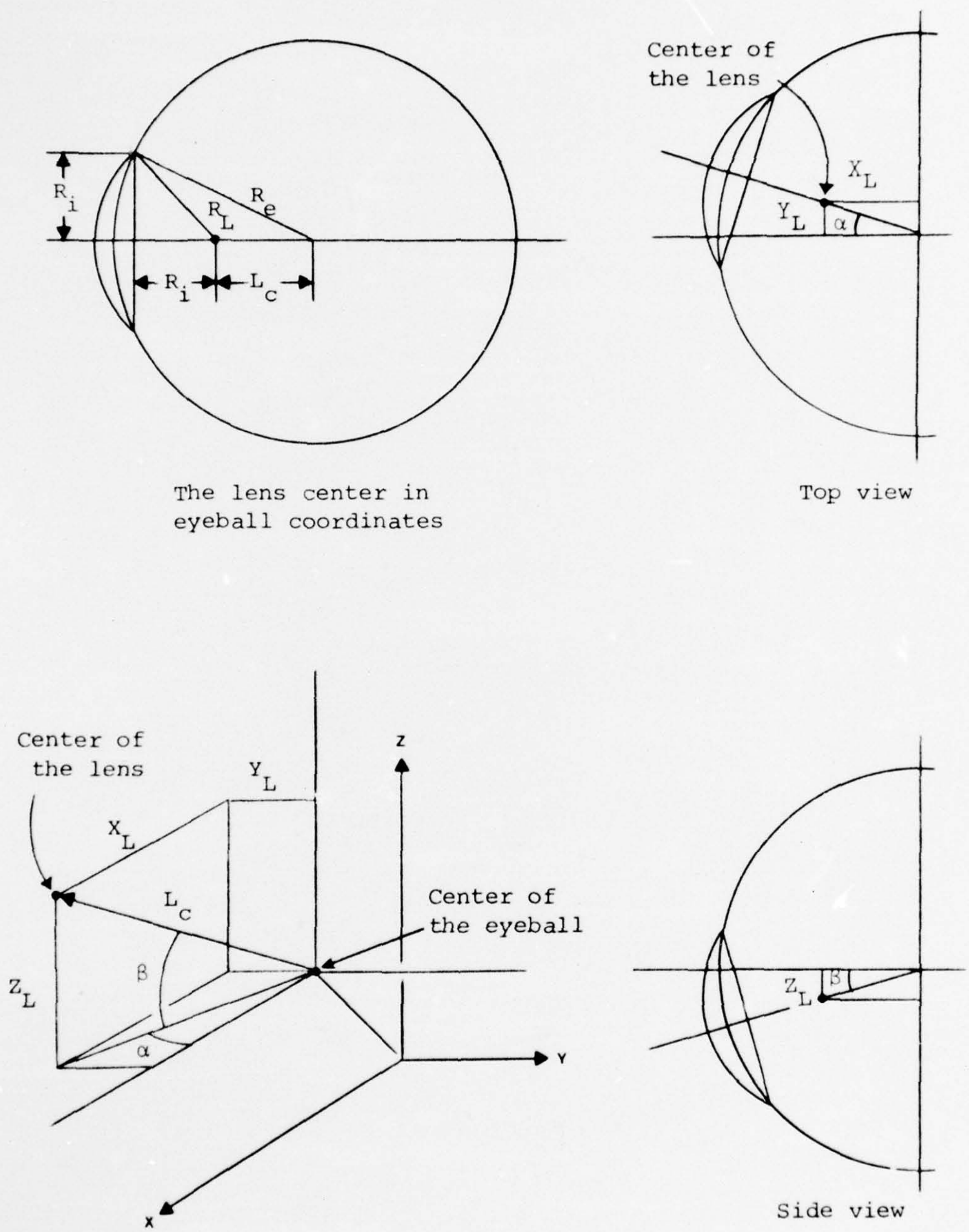


Figure A.2 - Locating the center of the lens sphere.

$$Z_L = L_c \sin(-\beta)$$
$$Y_L = L_c \cos(-\beta) \sin(\alpha)$$
$$X_L = L_c \cos(-\beta) \cos(\alpha)$$

The next step is to determine the reflection spot orientation angles θ and ϕ for each eye. When the reflection spot polygon is generated by the eyeball procedure it is centered on and perpendicular to the X axis of the eye. It is a distance R_L away from the center of the eye along the X axis. The angles θ and ϕ are used to rotate the spot polygon about the center of the eyeball into its proper orientation. The spot will then be displaced along the X axis a distance L_c. These rotations and displacement will place the reflection spot at the correct position on the surface of the eye lens. Refering to Figure A.3, the angles θ and ϕ are computed as follows.

$$\gamma_{Lgt} = \arctan\left((y_{Lgt}-y_{Lc})/(x_{Lgt}-x_{Lc})\right)$$
$$\gamma_V = \arctan\left((y_V-y_{Lc})/(x_V-x_{Lc})\right)$$
$$\gamma = (\gamma_{Lgt}+\gamma_V)/2$$
$$\theta = \gamma-\alpha$$
$$L_{Lgt} = \left((y_{Lgt}-y_{Lc})^2+(x_{Lgt}-x_{Lc})^2\right)^{1/2}$$
$$L_V = \left((y_V-y_{Lc})^2+(X_V-x_{Lc})^2\right)^{1/2}$$
$$\delta_{Lgt} = \arctan\left((z_{Lc}-z_{Lgt})/L_{Lgt}\right)$$
$$\delta_V = \arctan\left((z_{Lc}-z_V)/L_V\right)$$

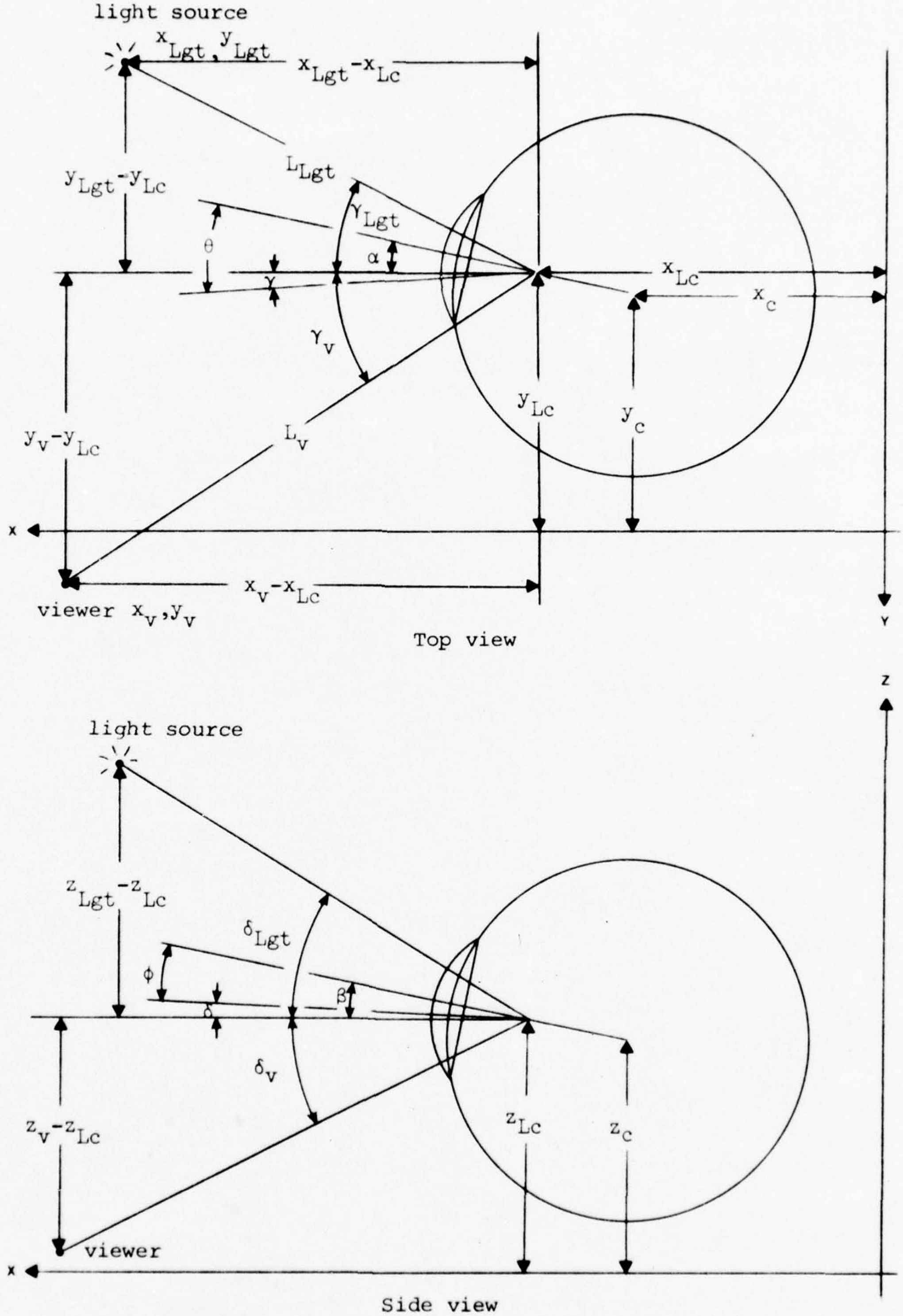


Figure A.3 - Reflection spot orientation angles

$$\delta = (\delta_{Lgt} + \delta_{V})/2$$

$$\phi = \delta - \beta$$

Note that in Figures A.1, A.2 and A.3 all angles are positive in the counter-clockwise sense.

Having determined all the orientation angles and displacements, how are they applied? For each eye the order of application is as follows. First the reflection spot polygon is rotated by θ and then by ϕ about the center of the eyeball. It is then displaced along the X axis of the eyeball by a distance L_c. Expressed in terms of transformation matrices this is

$$[T_{spot}] = [T_{\theta}][T_{\phi}][T_{L_c}]$$

Next the reflection spot and the rest of the eye are rotated about the center of the eyeball first by α and then by β.

$$[T_{orient}] = [T_{\alpha}][T_{\beta}]$$

And finally the entire eye is displaced by x_c, y_c, z_c to assume its proper position in the face.

$$[T_{eye}] = [T_{orient}][T_{face}]$$

The reflection spot is transformed first by $[T_{spot}]$ and then by $[T_{eye}]$. The rest of the eyeball is transformed only by $[T_{eye}]$.

APPENDIX B

THE PARAMETER VALUES USED FOR THE SPEECH SYNCHRONIZED ANIMATED SEGMENTS

This appendix contains graphs of the parameter values used to generate six computer animated segments (see Appendix D). These segments were animated to match a spoken soundtrack. The soundtrack is the reading of a short poem. The poem, "Little Stone" by Emily Dickinson, is listed below.

How happy is the little stone
That rambles in the road alone,
And never cares about careers
And exigencies never fears;
Whose coat of elemental brown
A passing universe put on;
And independent as the sun,
Associates or glows alone,
Fulfilling absolute decree
In casual simplicity.

In the following figures the vertical axes correspond to the parameter values while the horizontal axes represent time expressed in terms of frame numbers. Also indicated on each graph is the phonetic representation [11] of the soundtrack.

The actual data values used in specifying the parameter functions are indicated as dots. The intermediate values were generated using the parameter change functions discussed in Chapter 4.

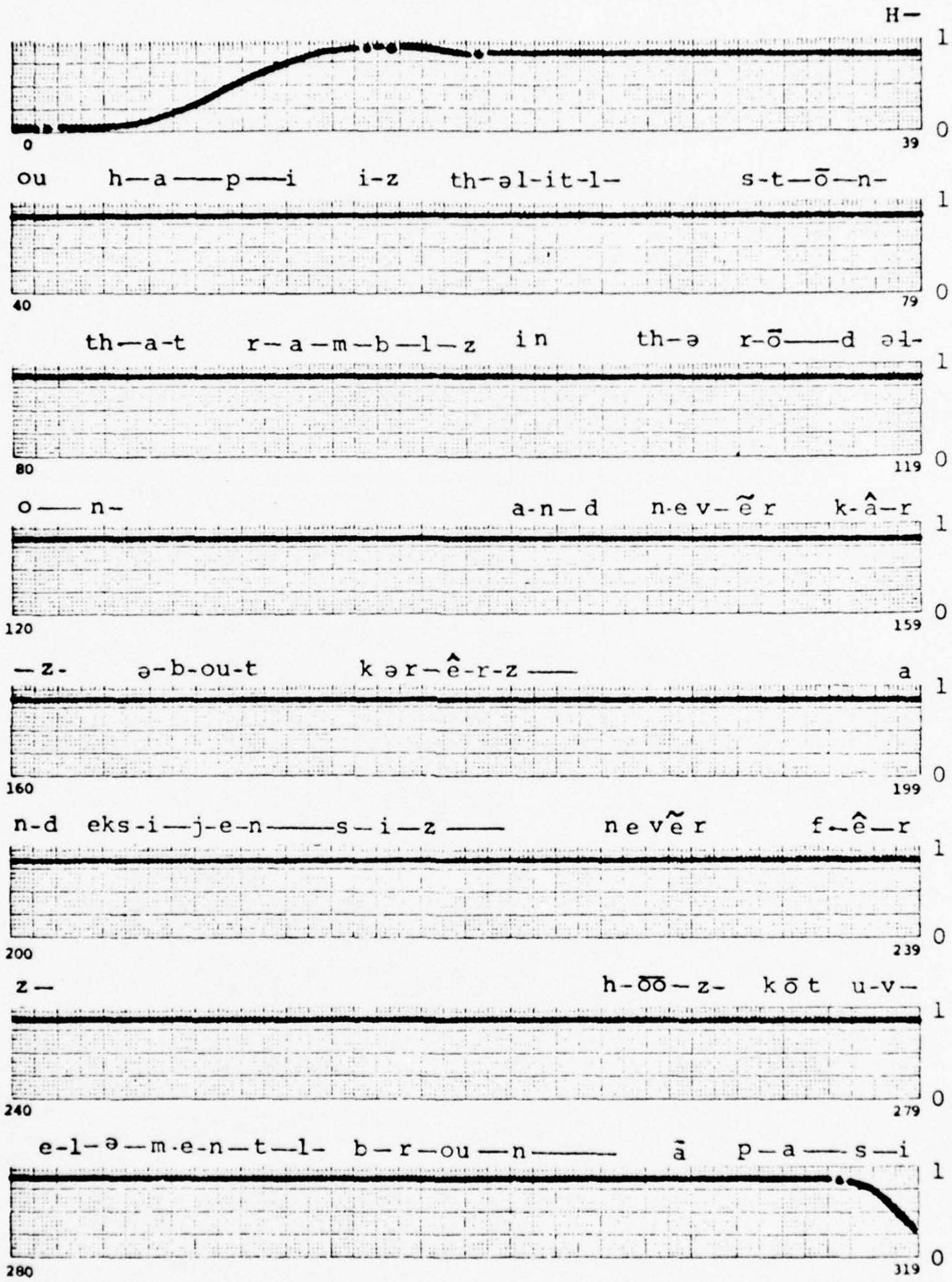


Figure B.1 - The values of the eyelid opening parameter for all six sequences. A value of zero corresponds to the eyelid being closed, a value of one corresponds to the eyelid being wide open.

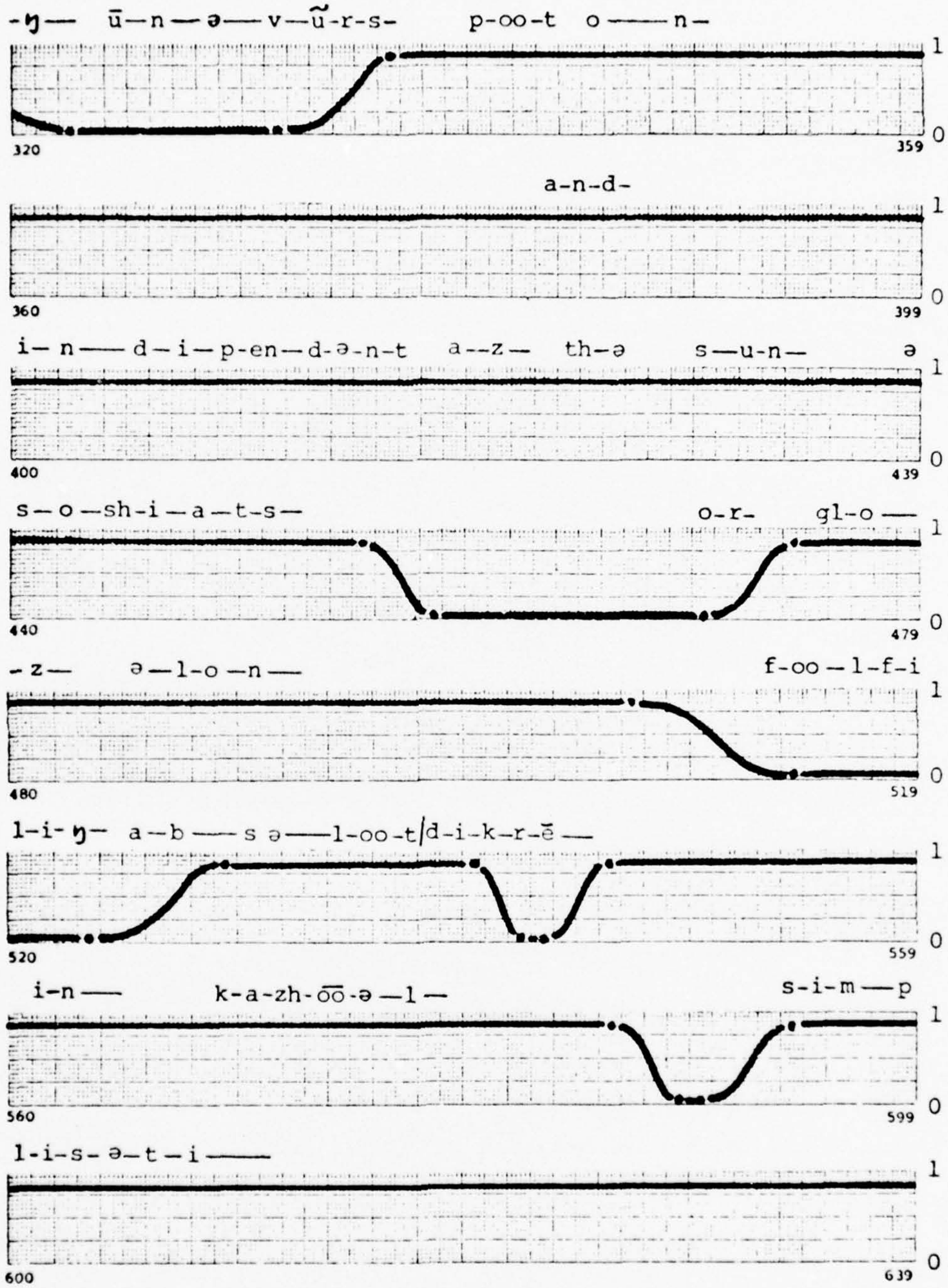


Figure B.1 (continued)

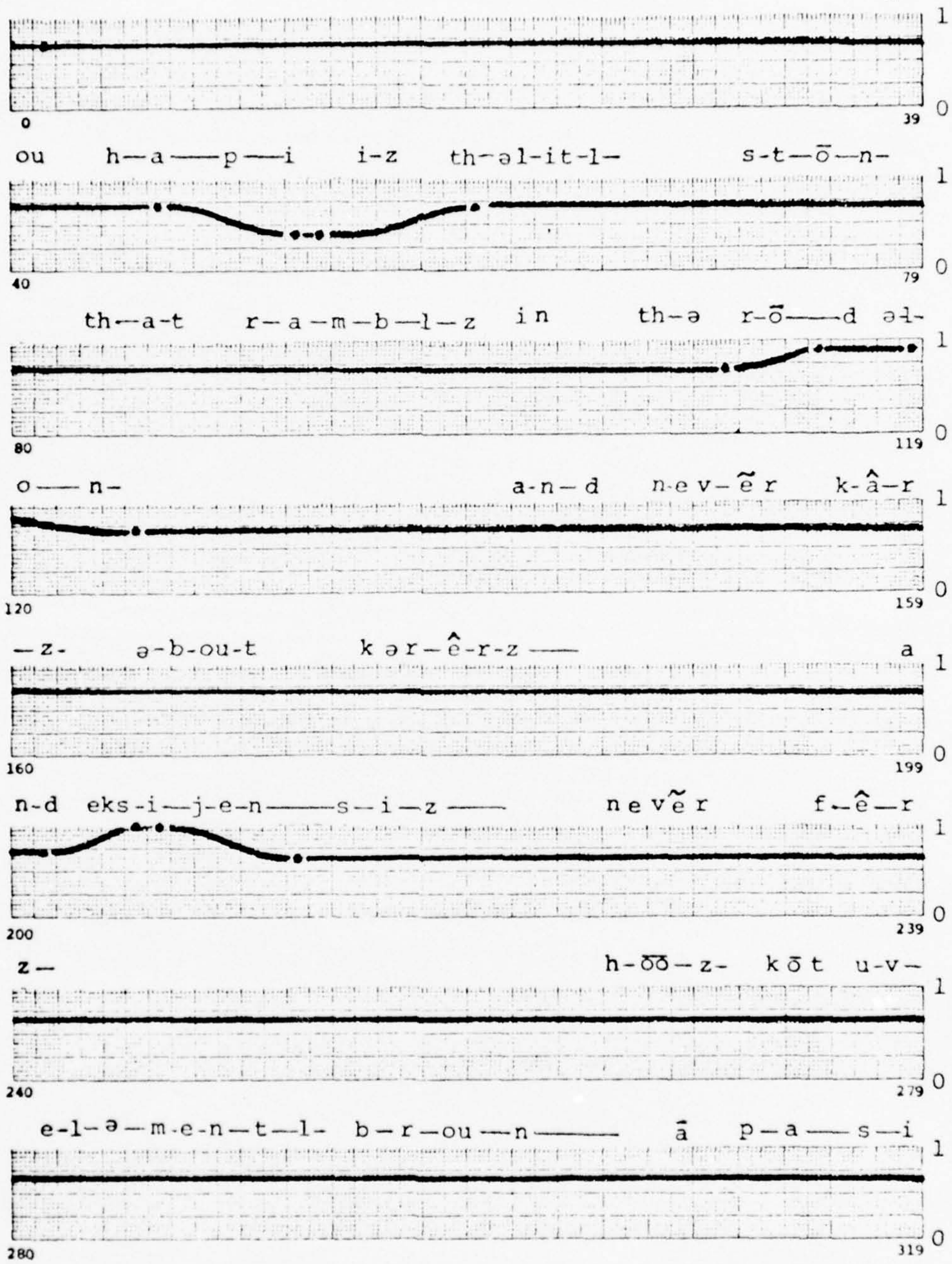


Figure B.2 - The values of the eyebrow arch parameter for sequences 1 through 5. A value of 0 corresponds to fully arched eyebrows. A value of 1 corresponds to minimum eyebrow arch. In frames 320 to 639 the parameter value is .65.

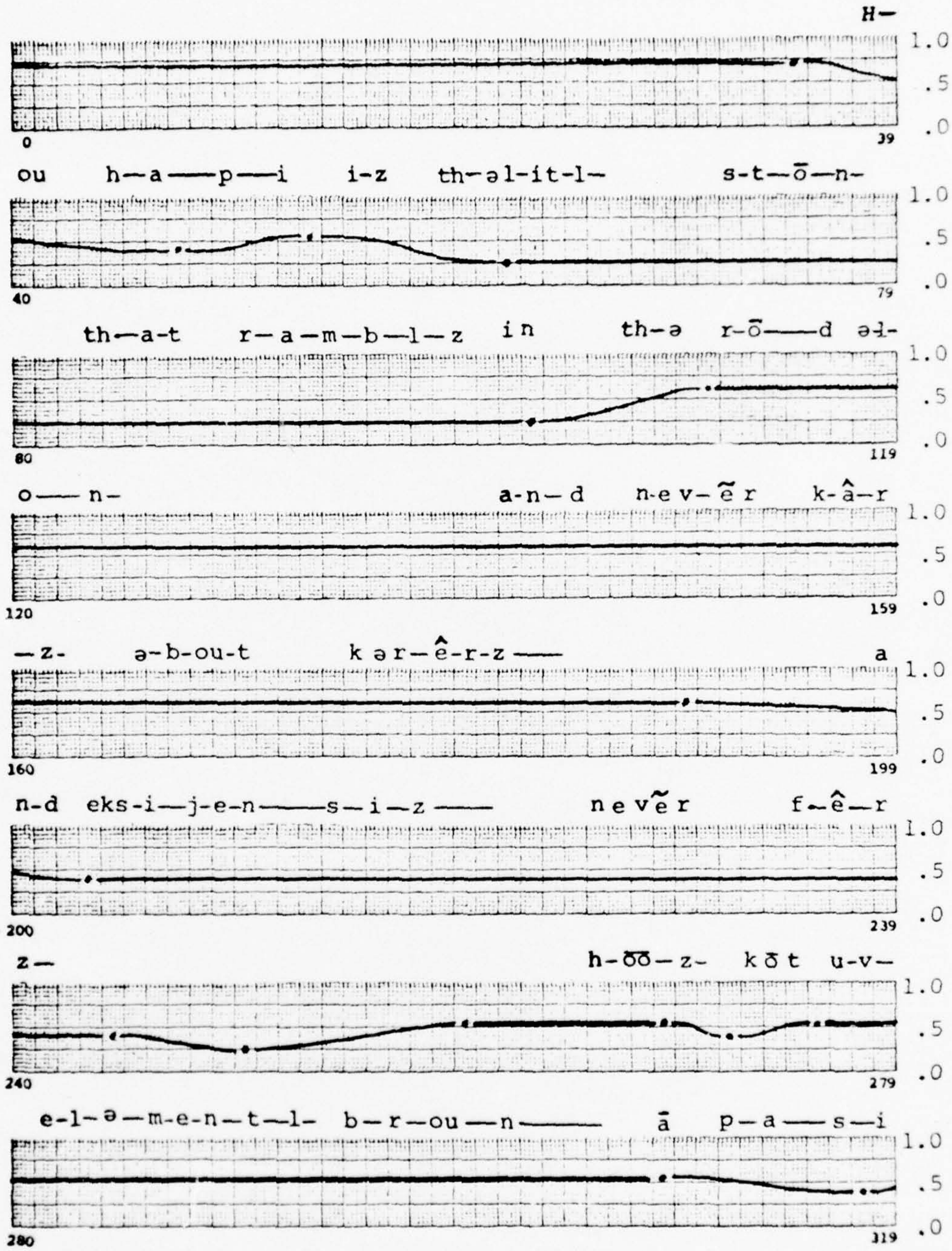


Figure B.3 - The values of the eyebrow arch parameter for sequence 6.

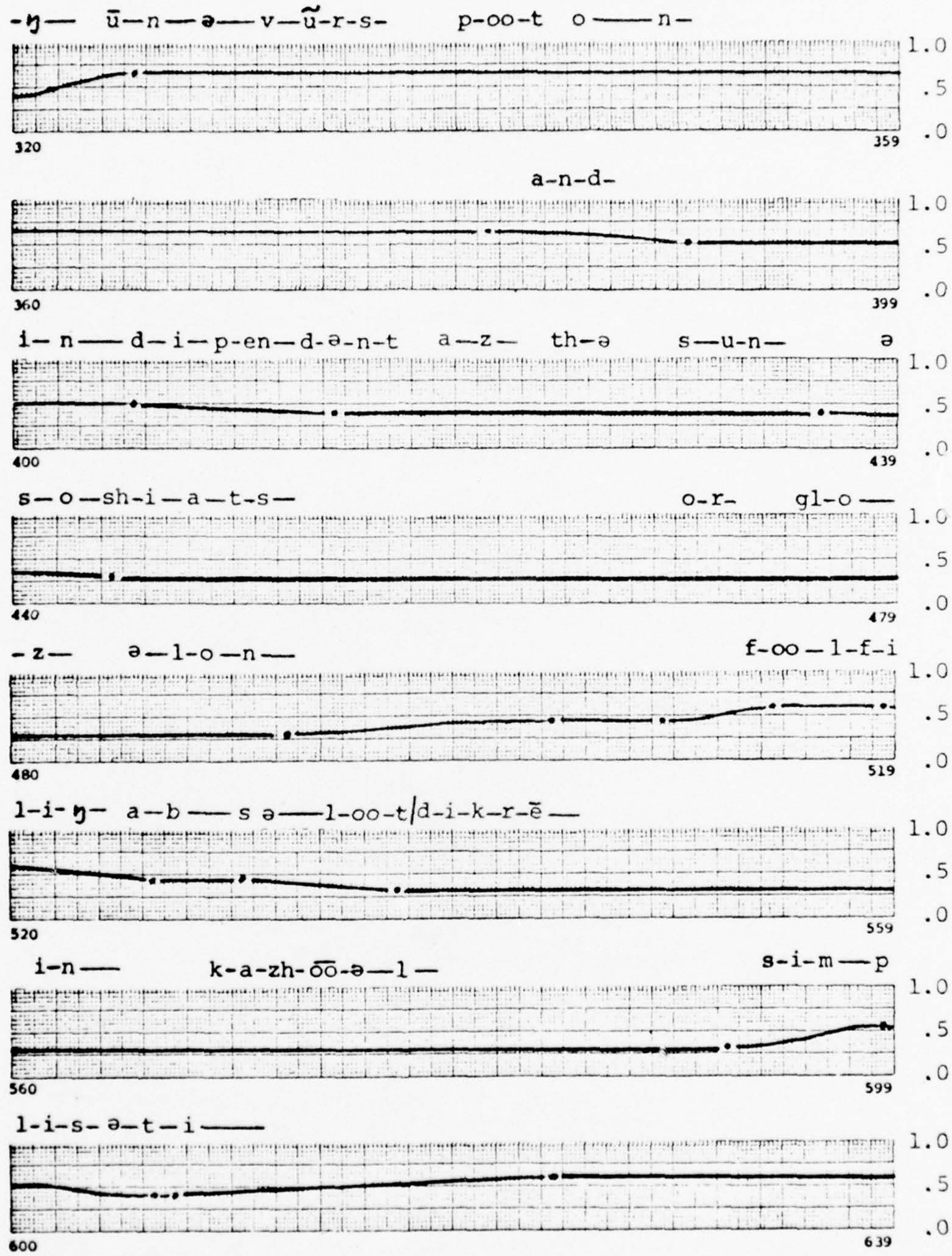


Figure B.3 (continued)

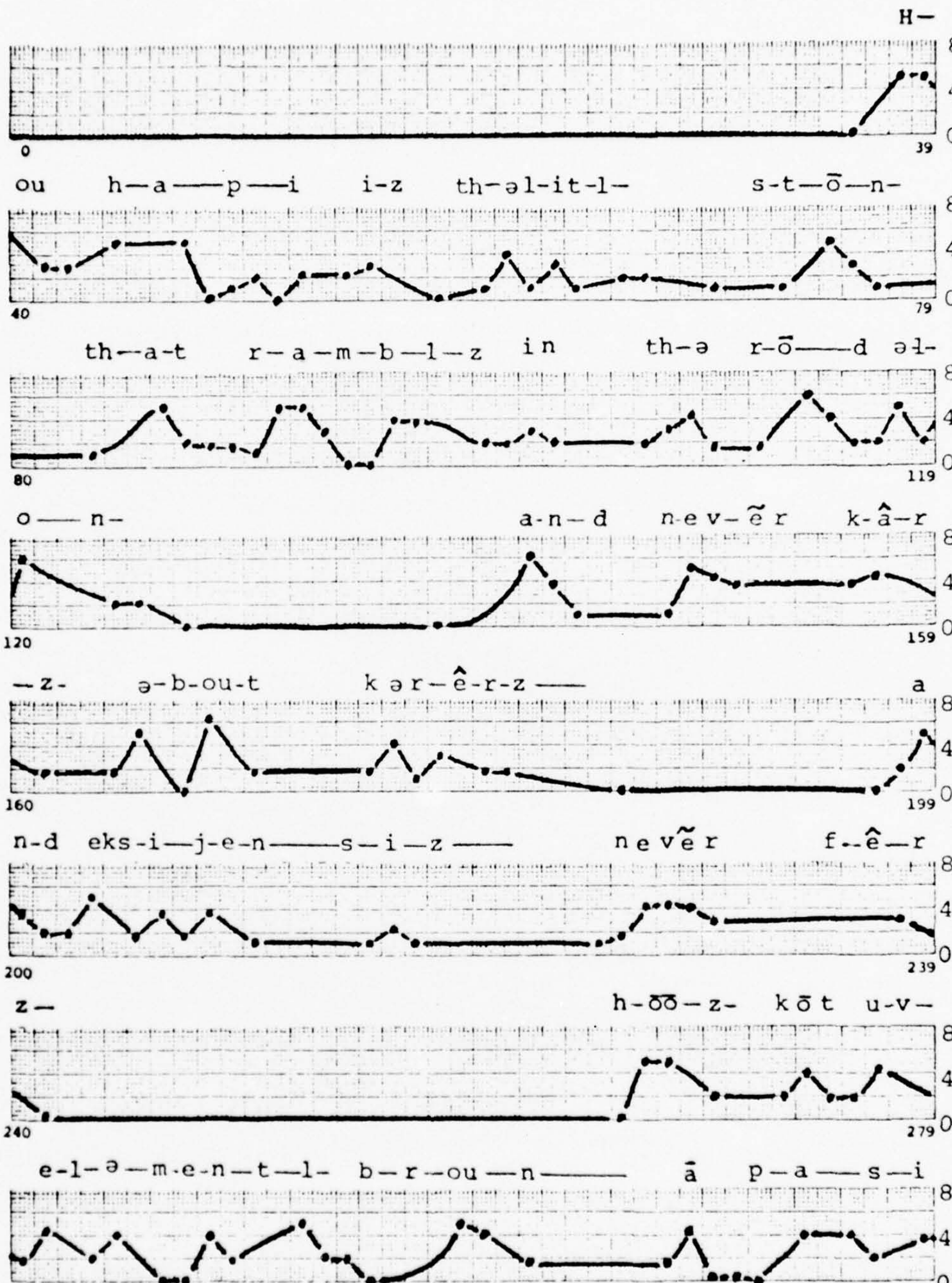


Figure B.4 - The jaw rotation values used for sequence 1.

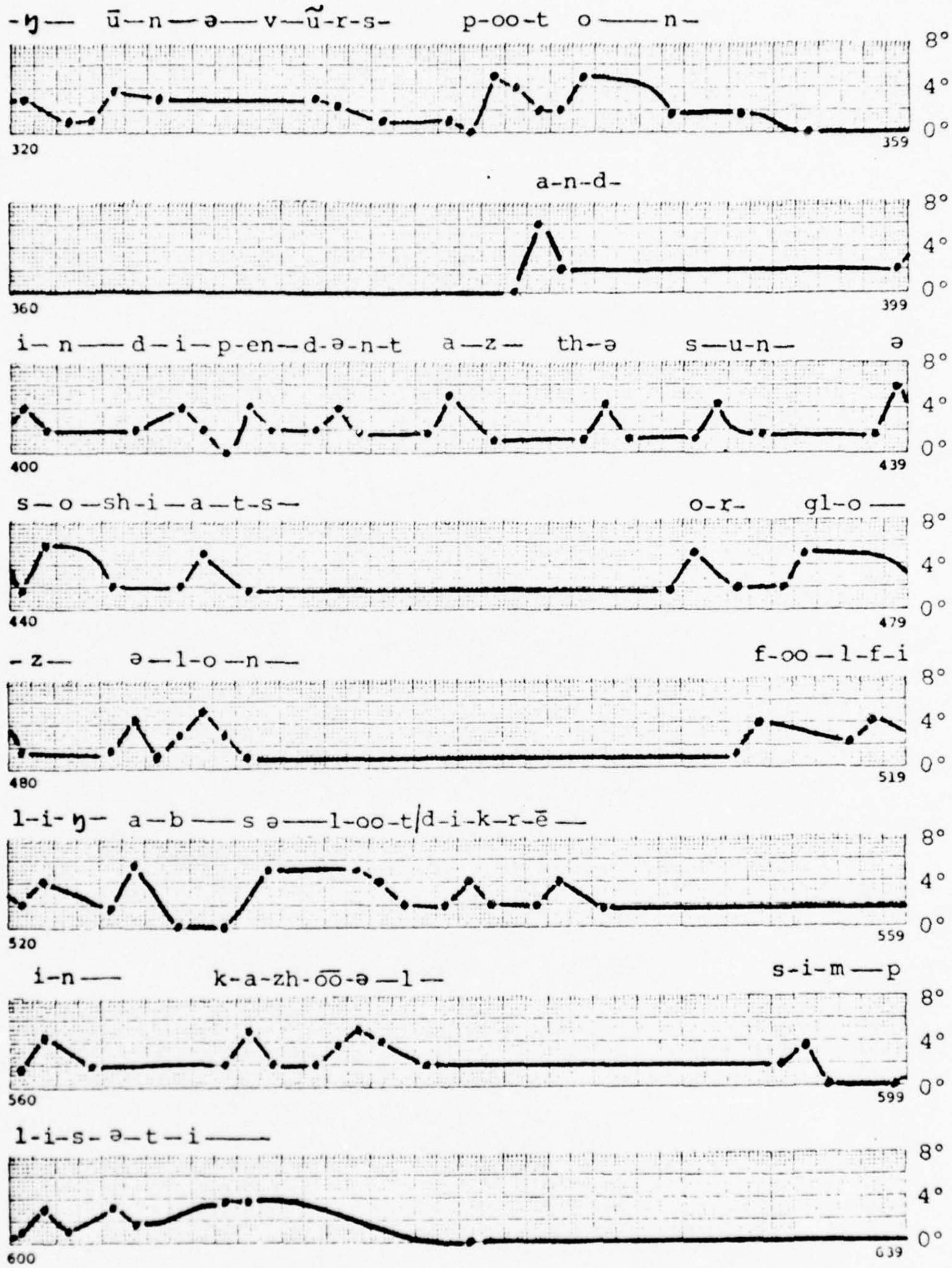


Figure B.4 (continued)

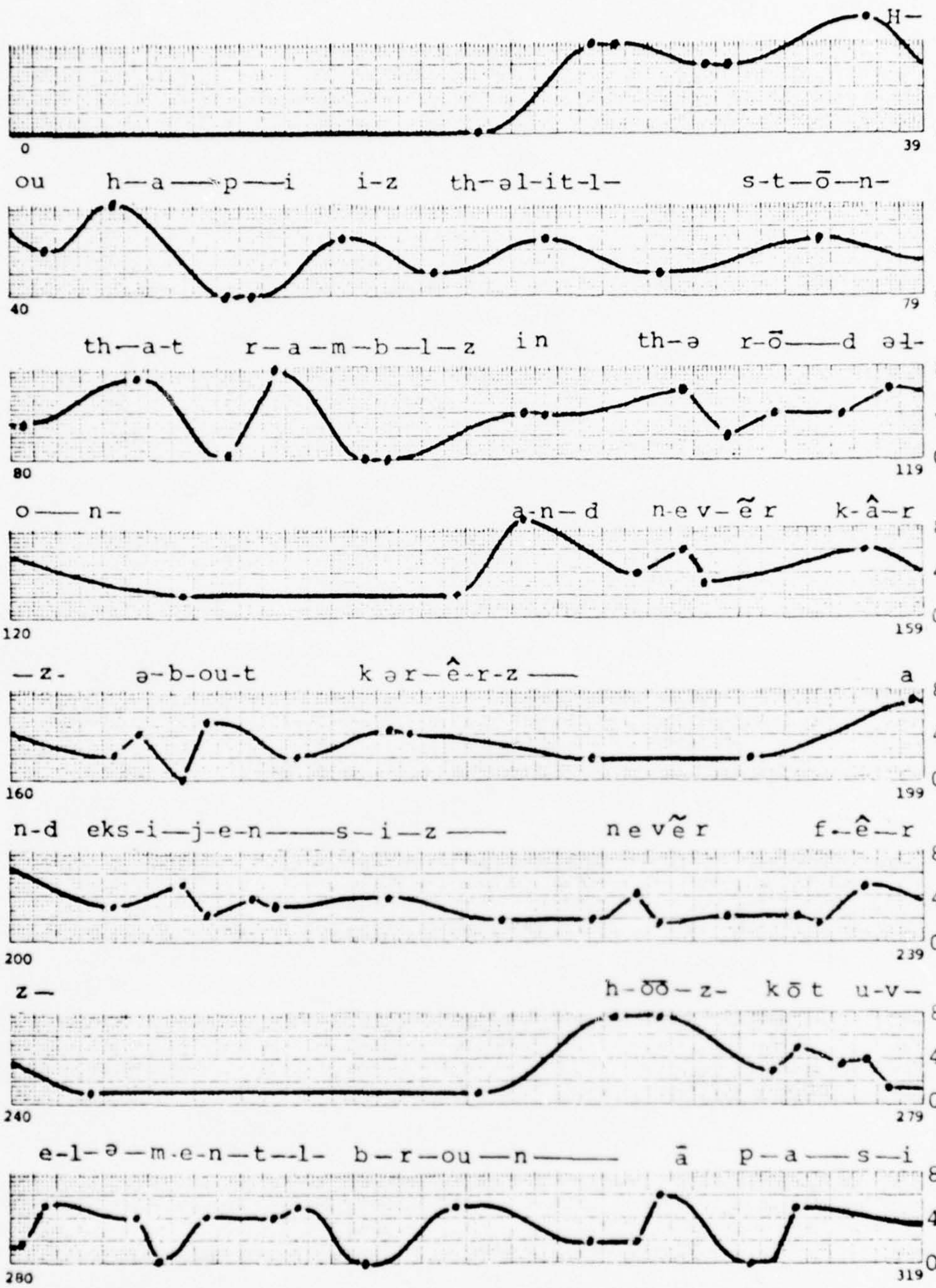


Figure B.5 - The jaw rotation values used for sequence 2.

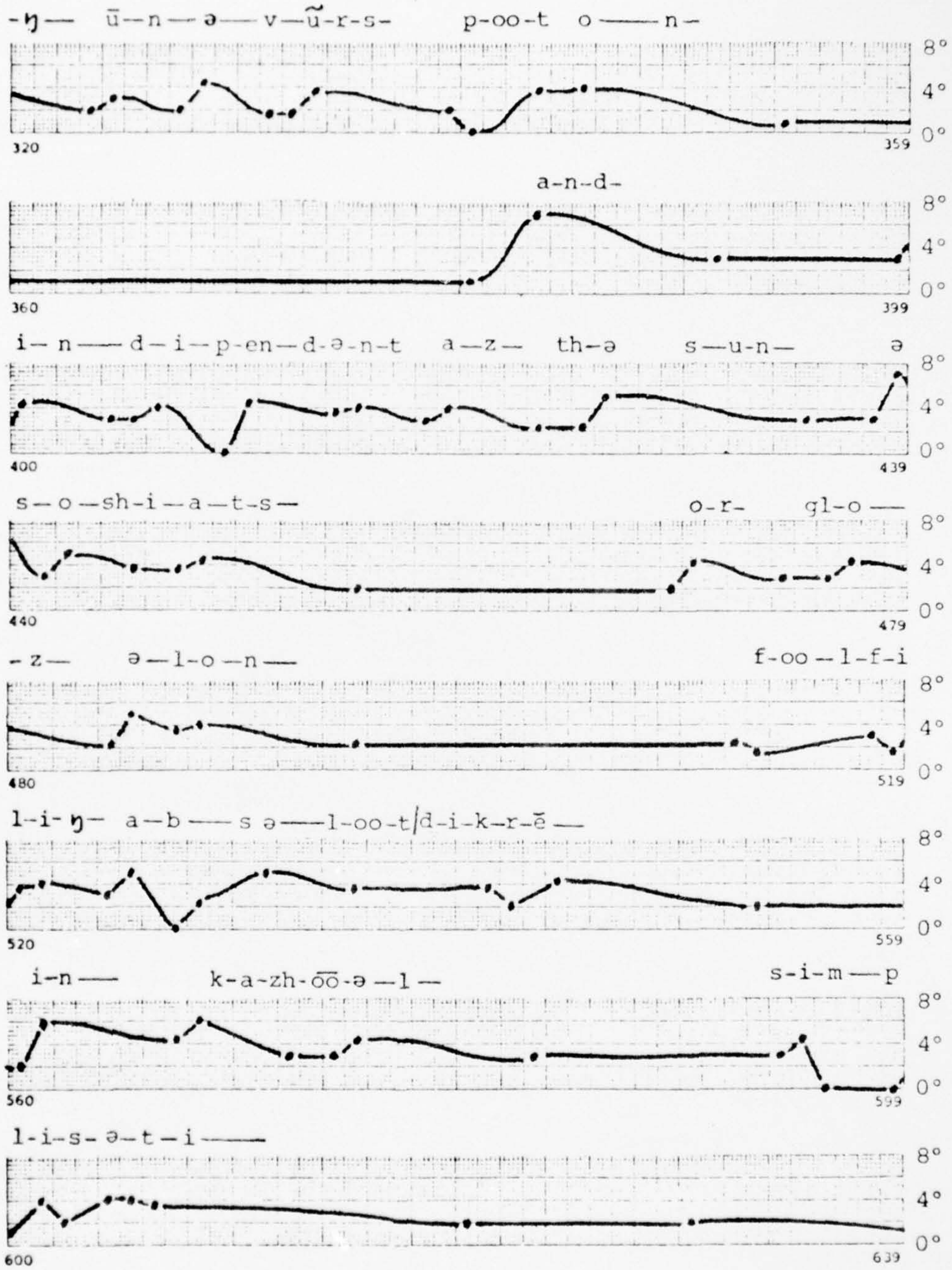


Figure B.5 (continued)

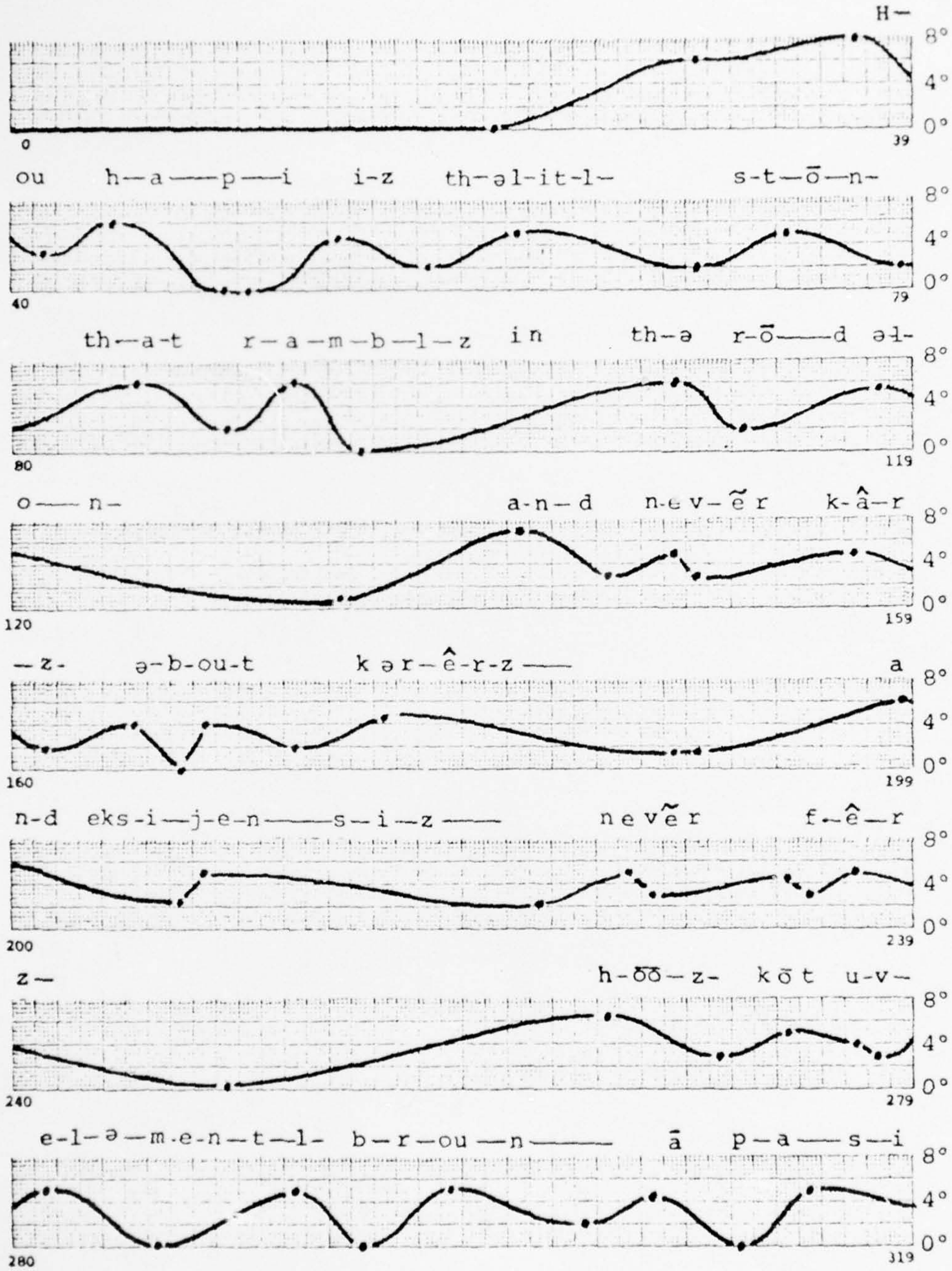


Figure B.6 - The jaw rotation values used for sequences 3 and 4.

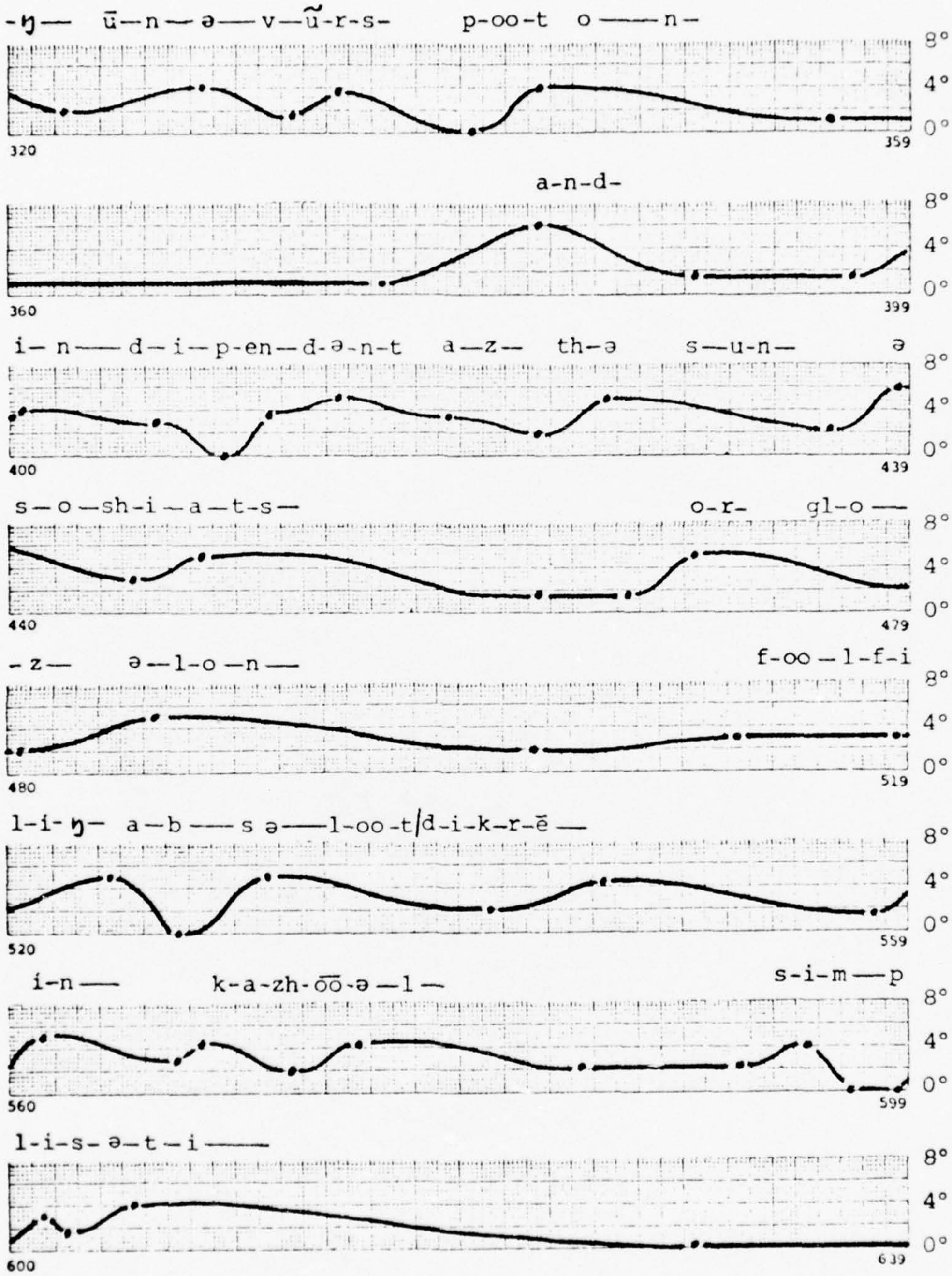


Figure B.6 (continued)

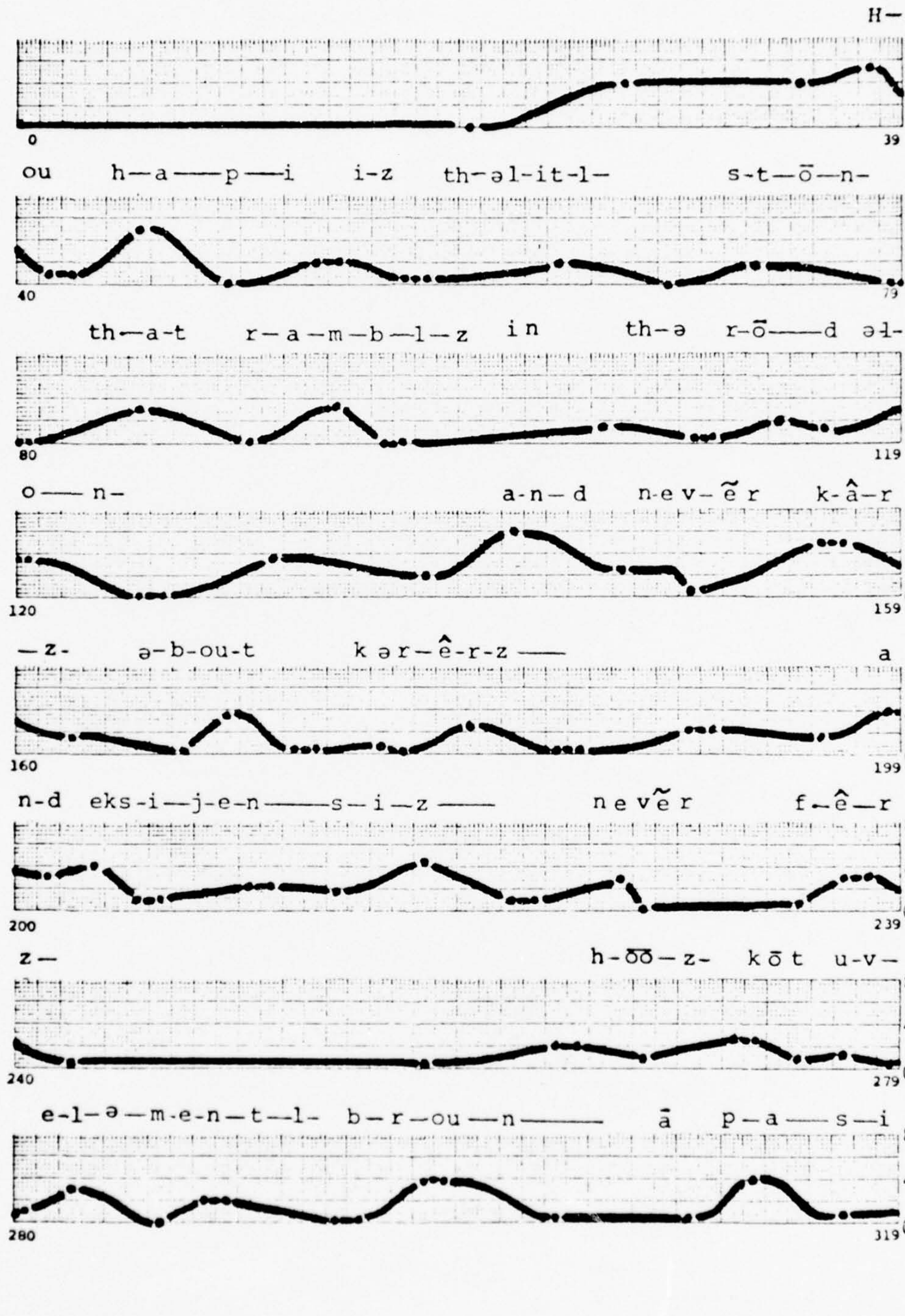


Figure B.7 - The jaw rotation values used for sequence 5.

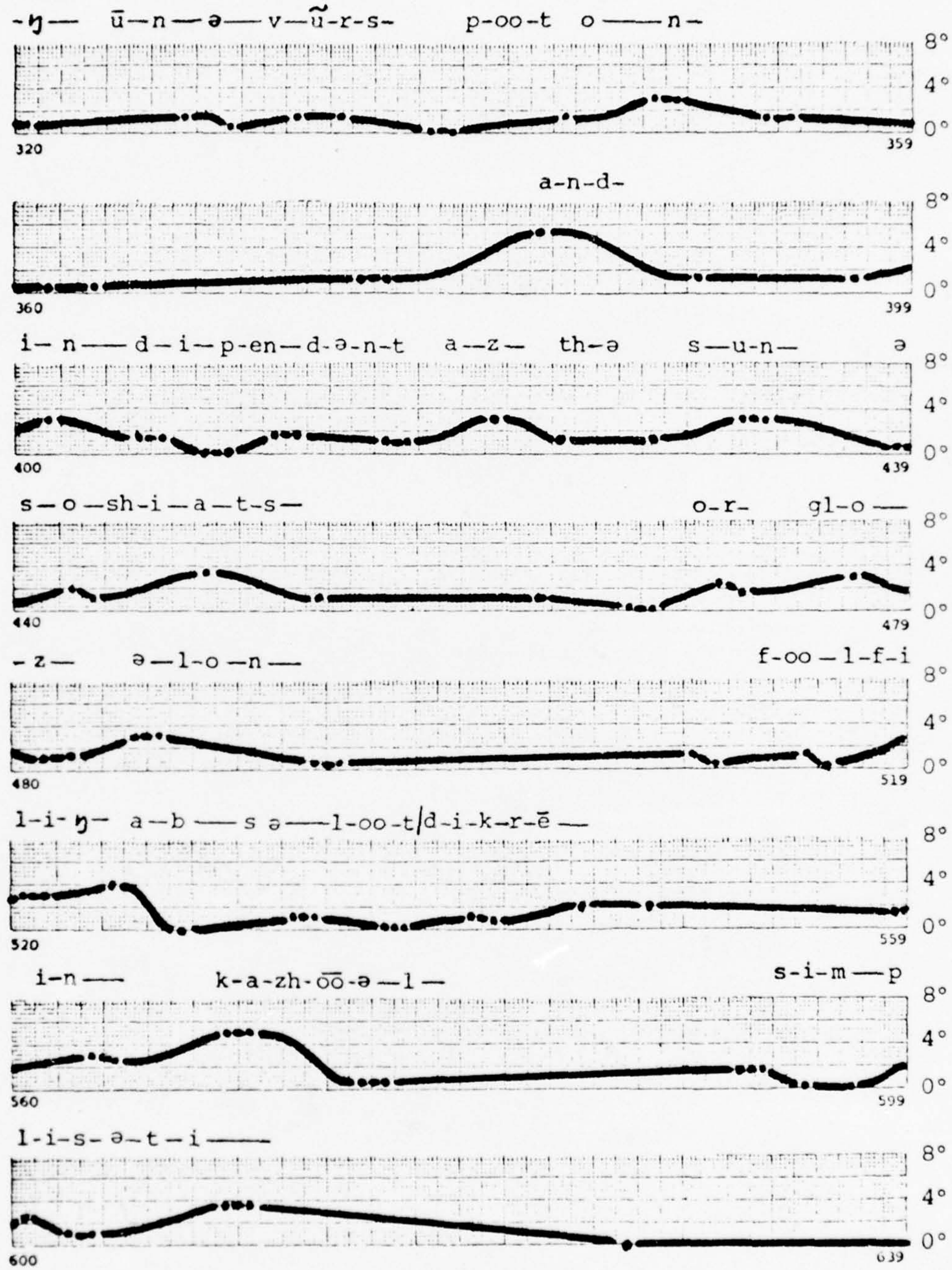


Figure B.7 (continued)

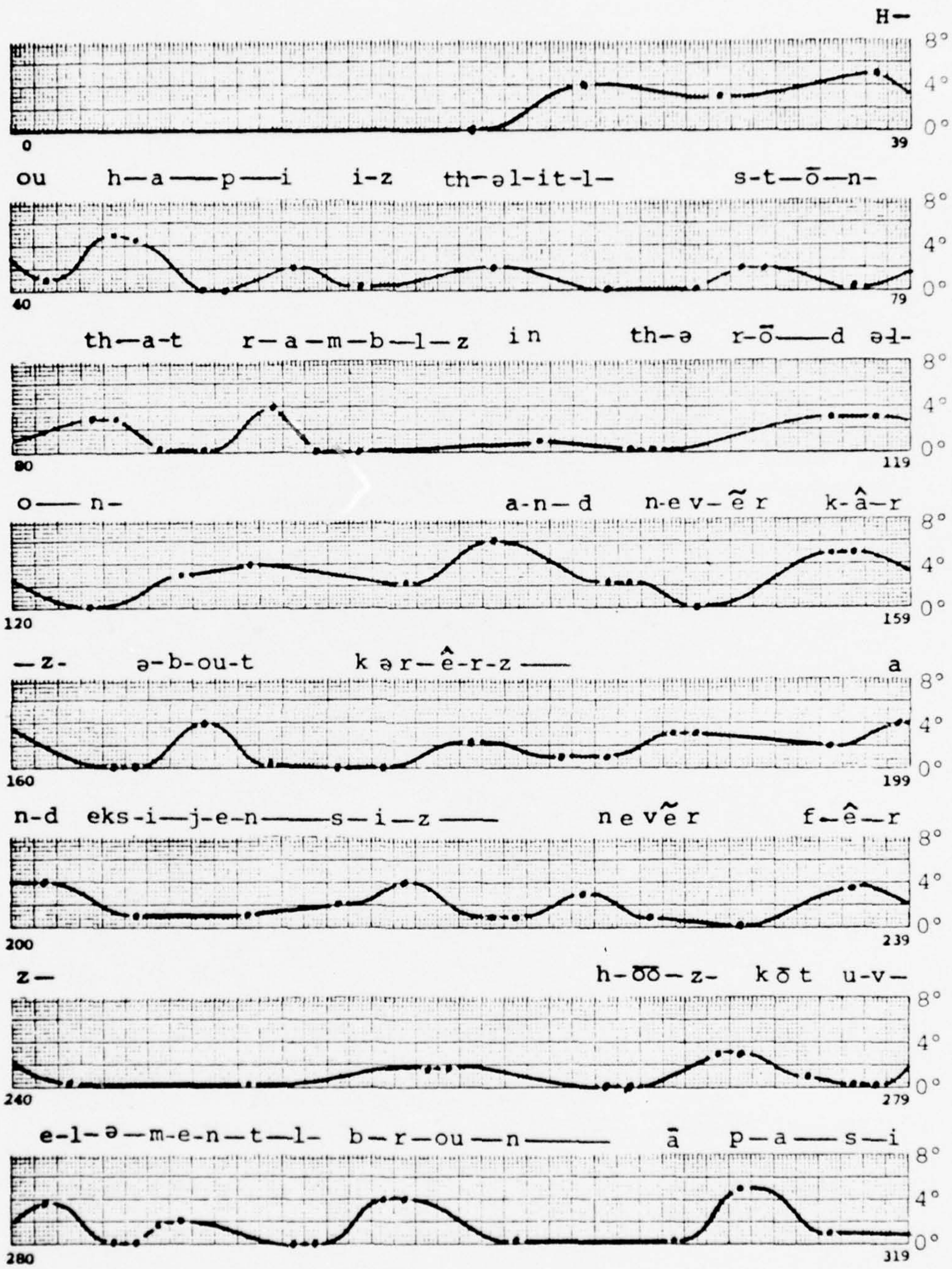


Figure B.8 - The jaw rotation values used for sequence 6.

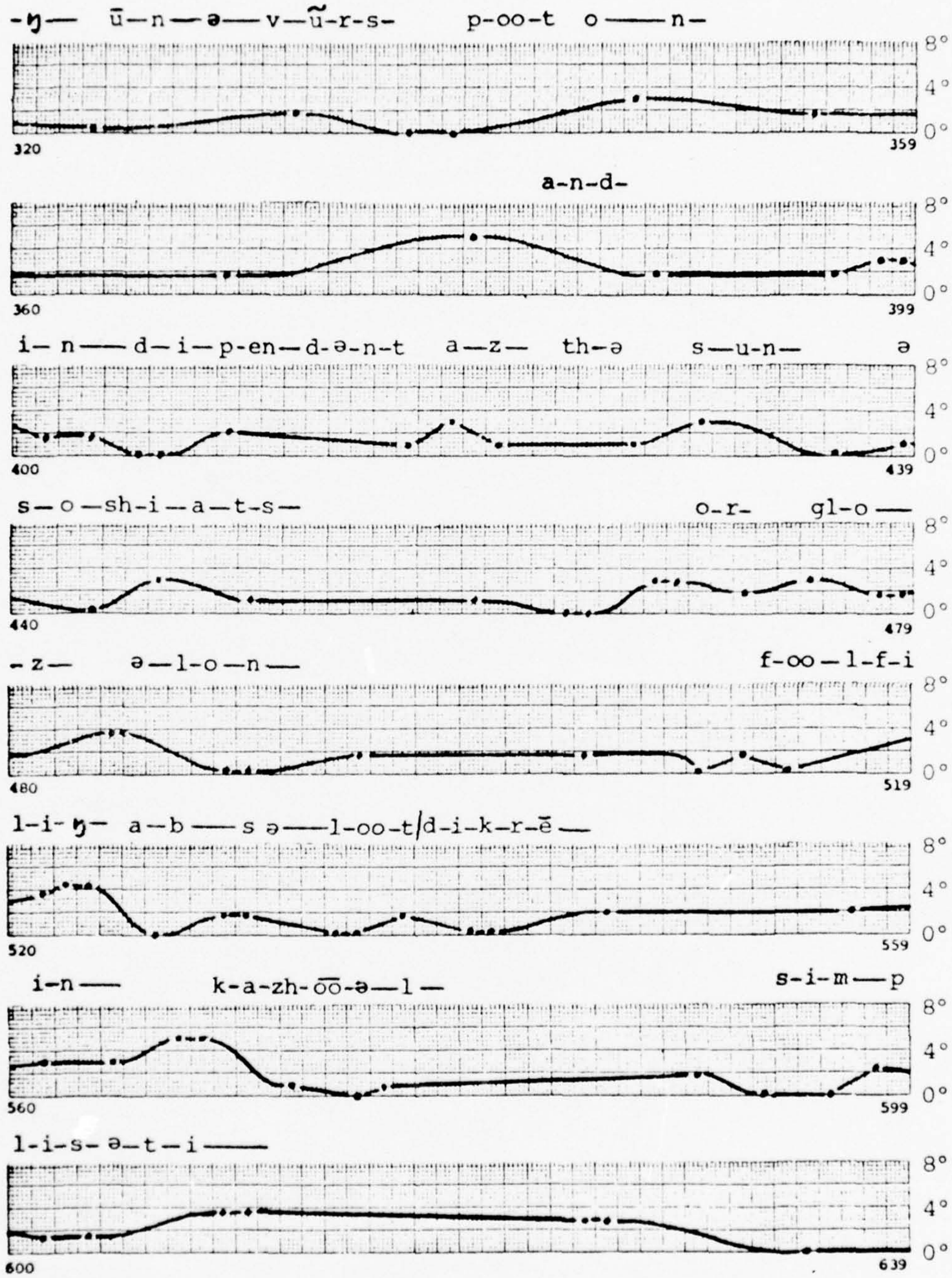


Figure B.8 (continued)

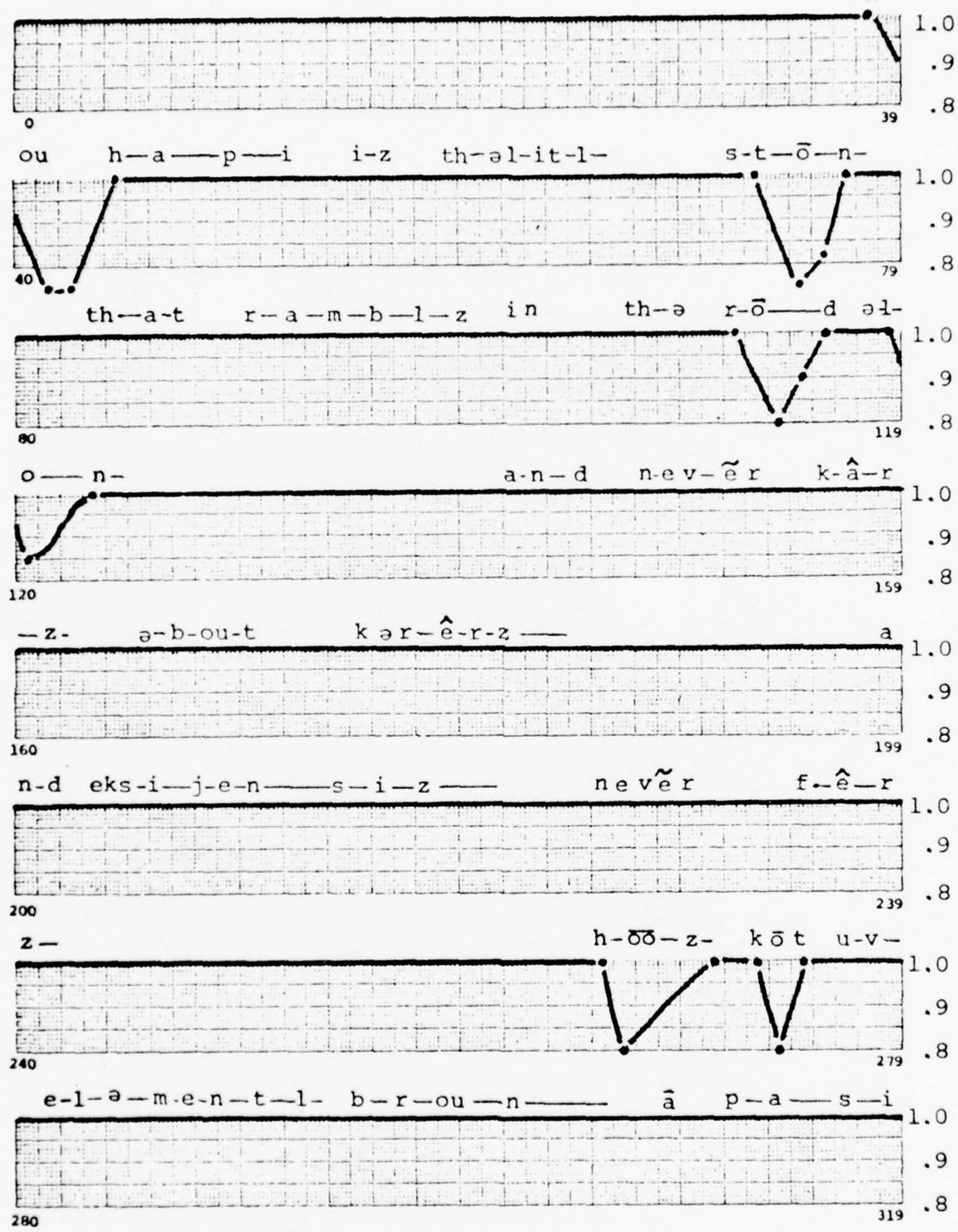


Figure B.9 - The scale factors used for mouth width in sequence 1.

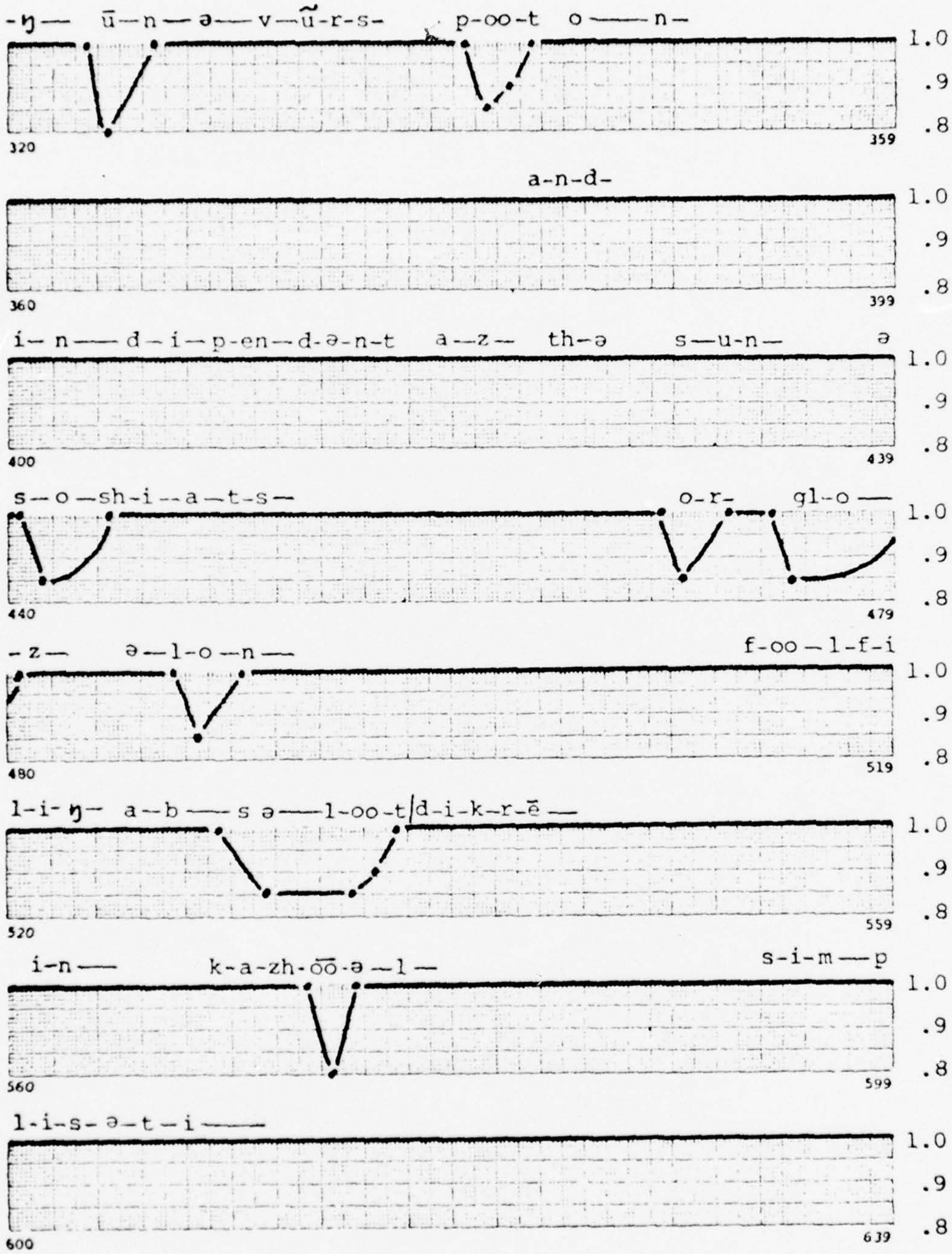


Figure B.9 (continued)

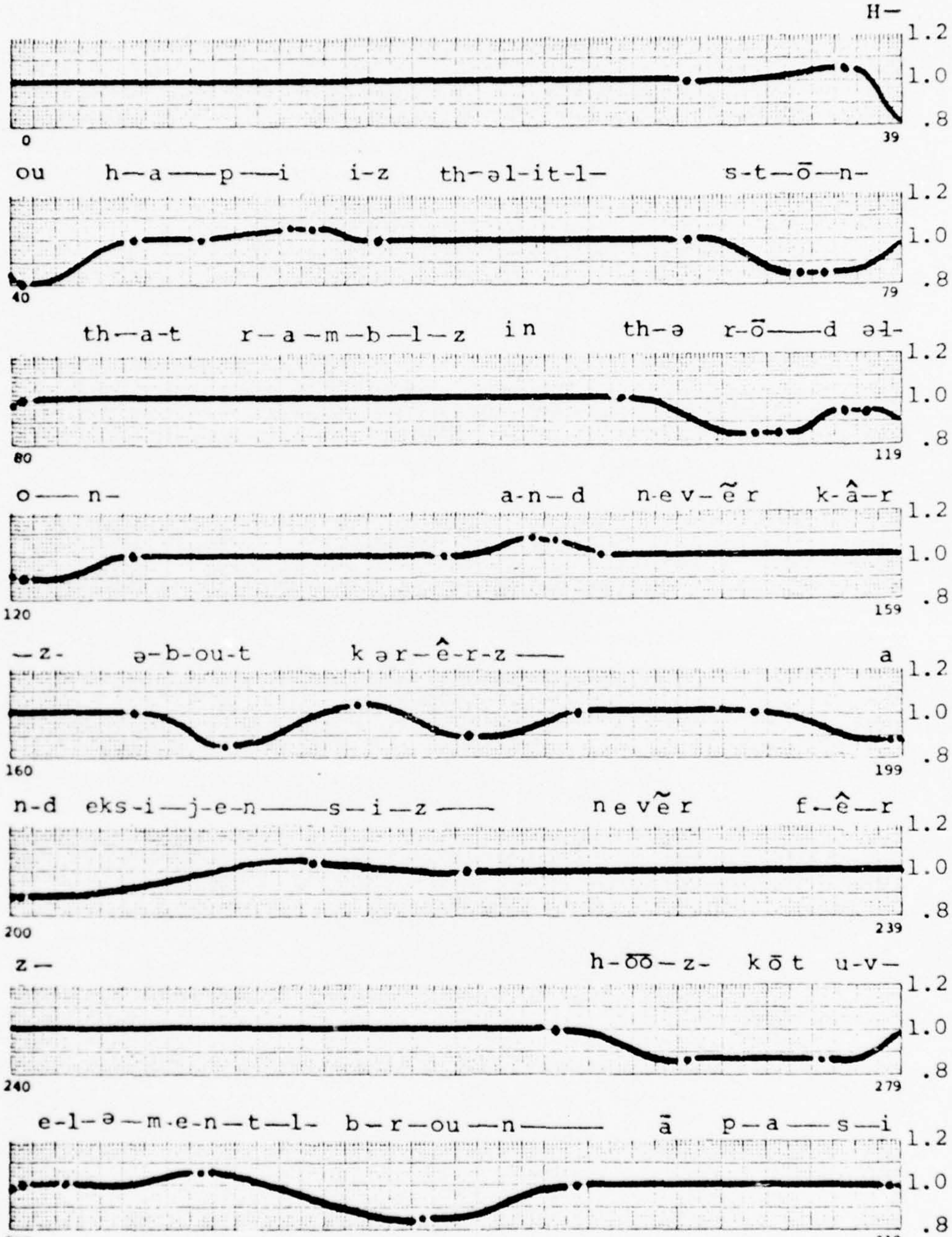


Figure B.10 - The scale factors used for mouth width in sequence 4.

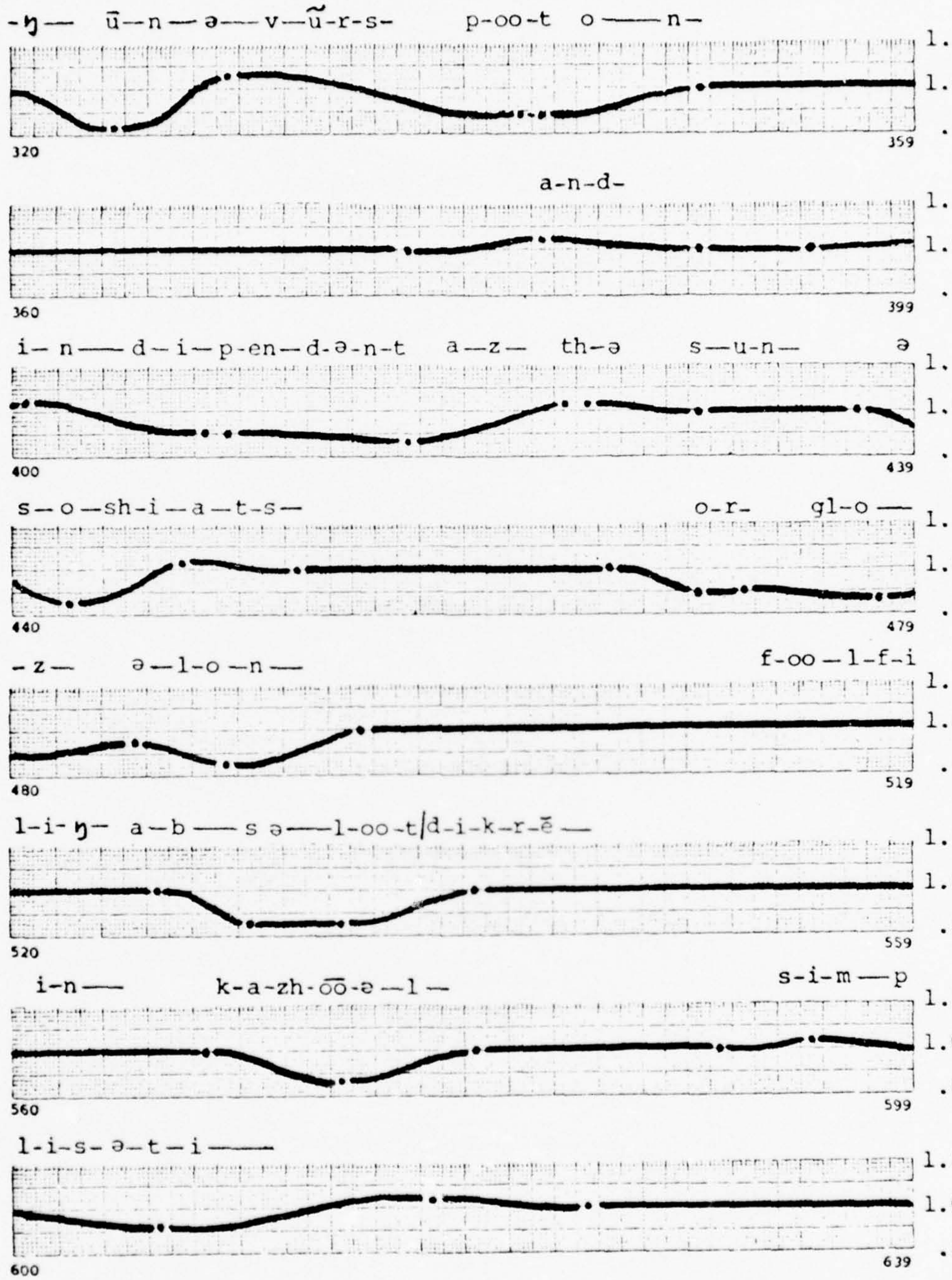


Figure B.10 (continued)

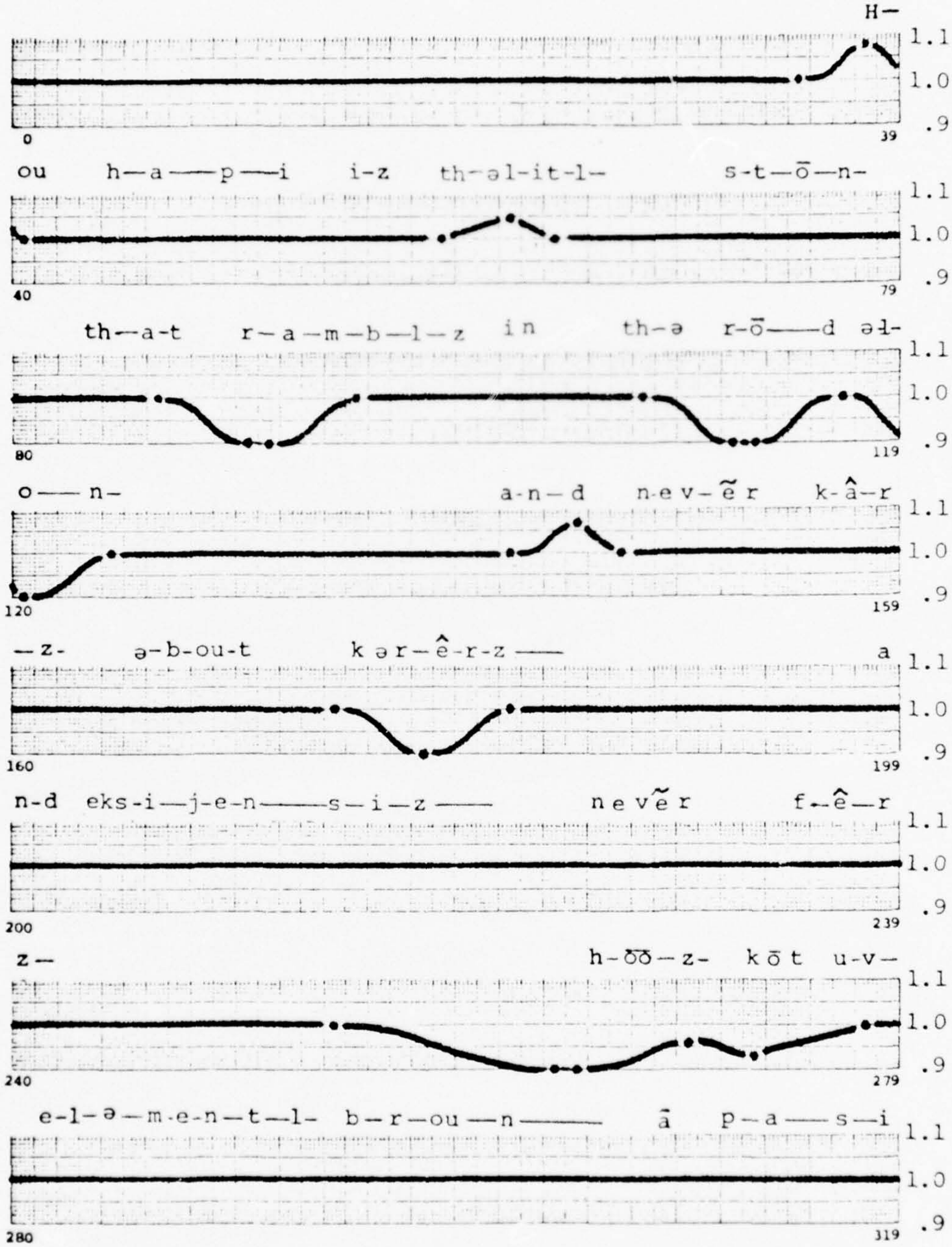


Figure B.11 - The scale factors used for mouth width in sequence 5.

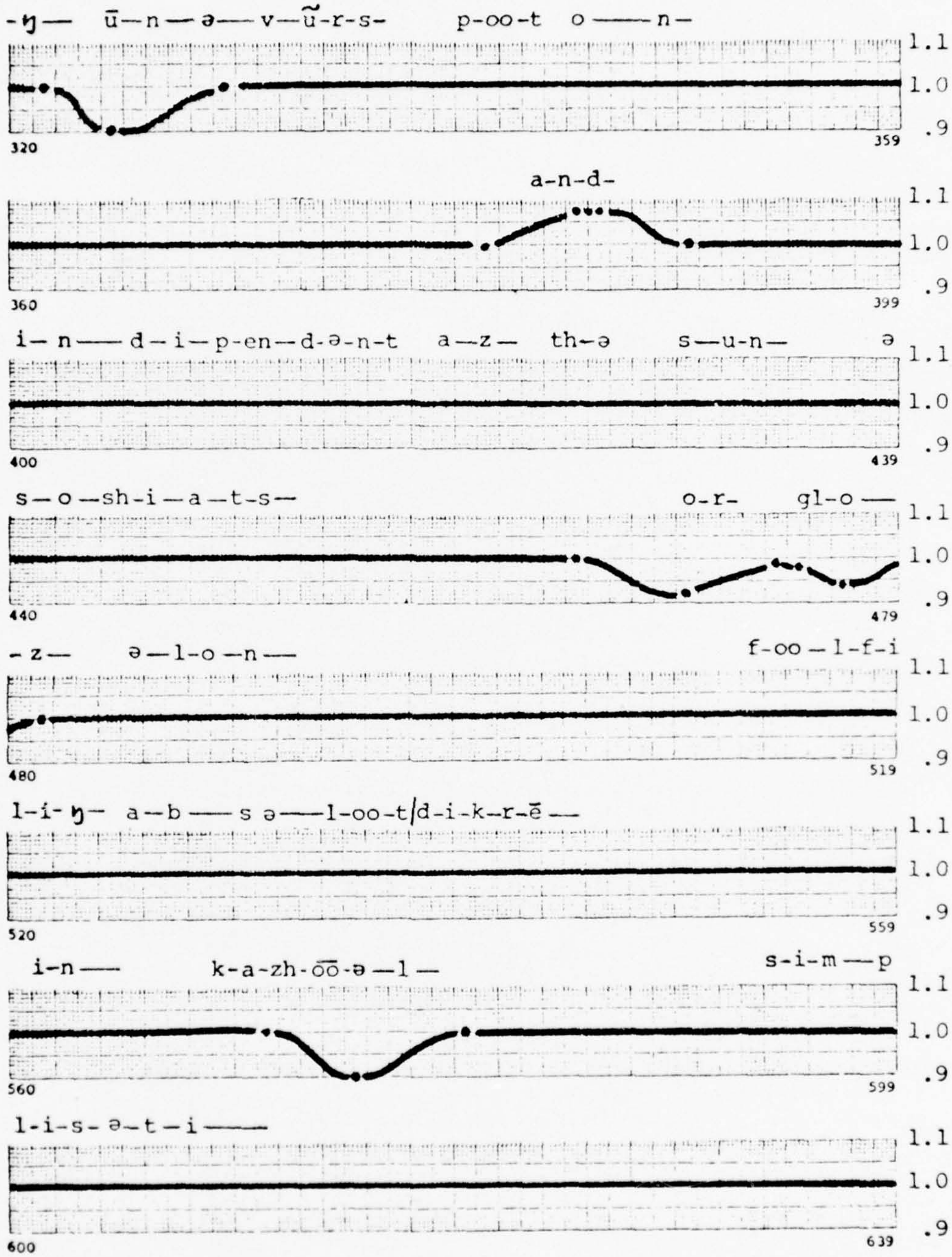


Figure B.11 (continued)

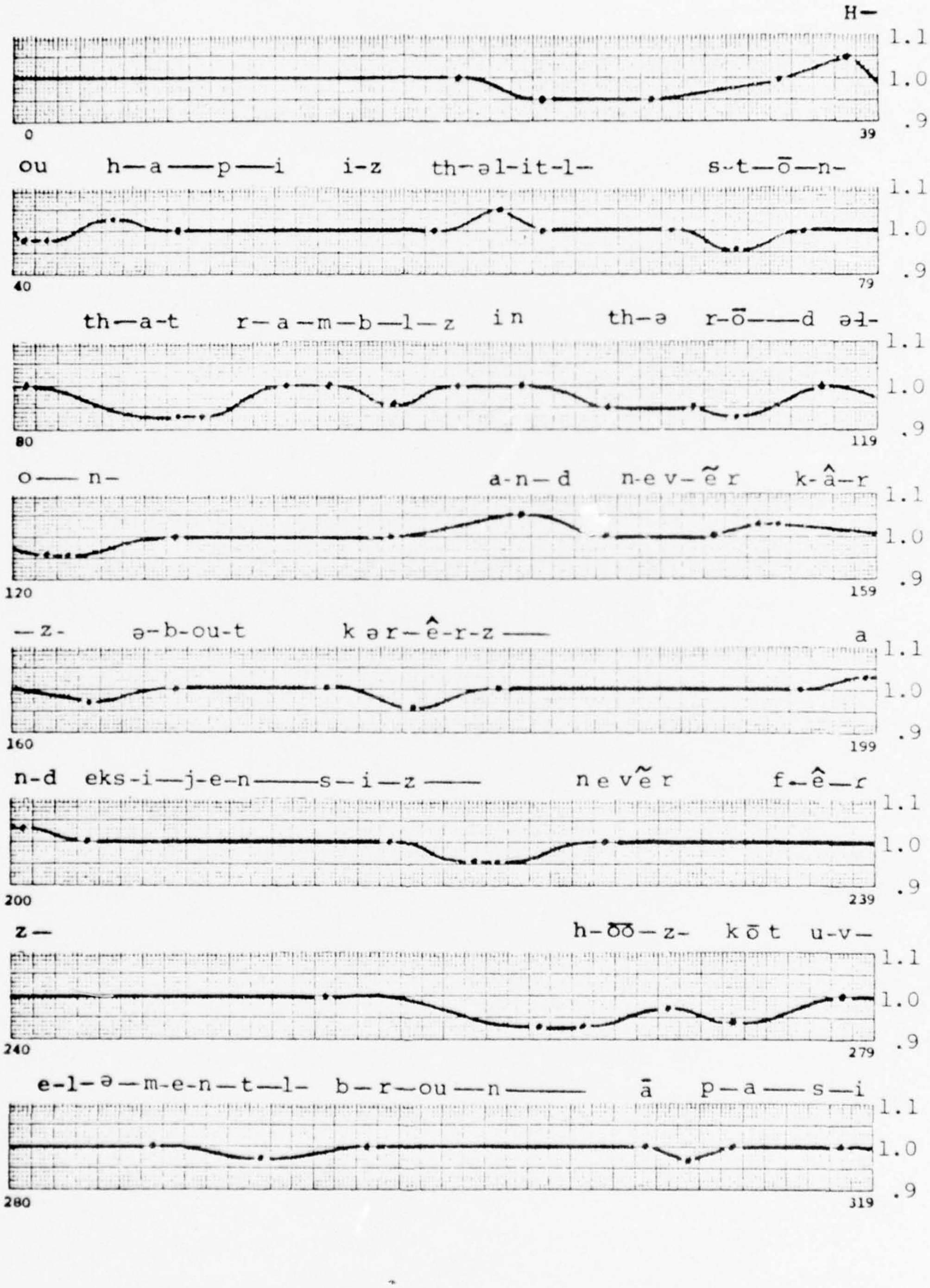


Figure B.12 - The scale factors used for mouth width in sequence 6.

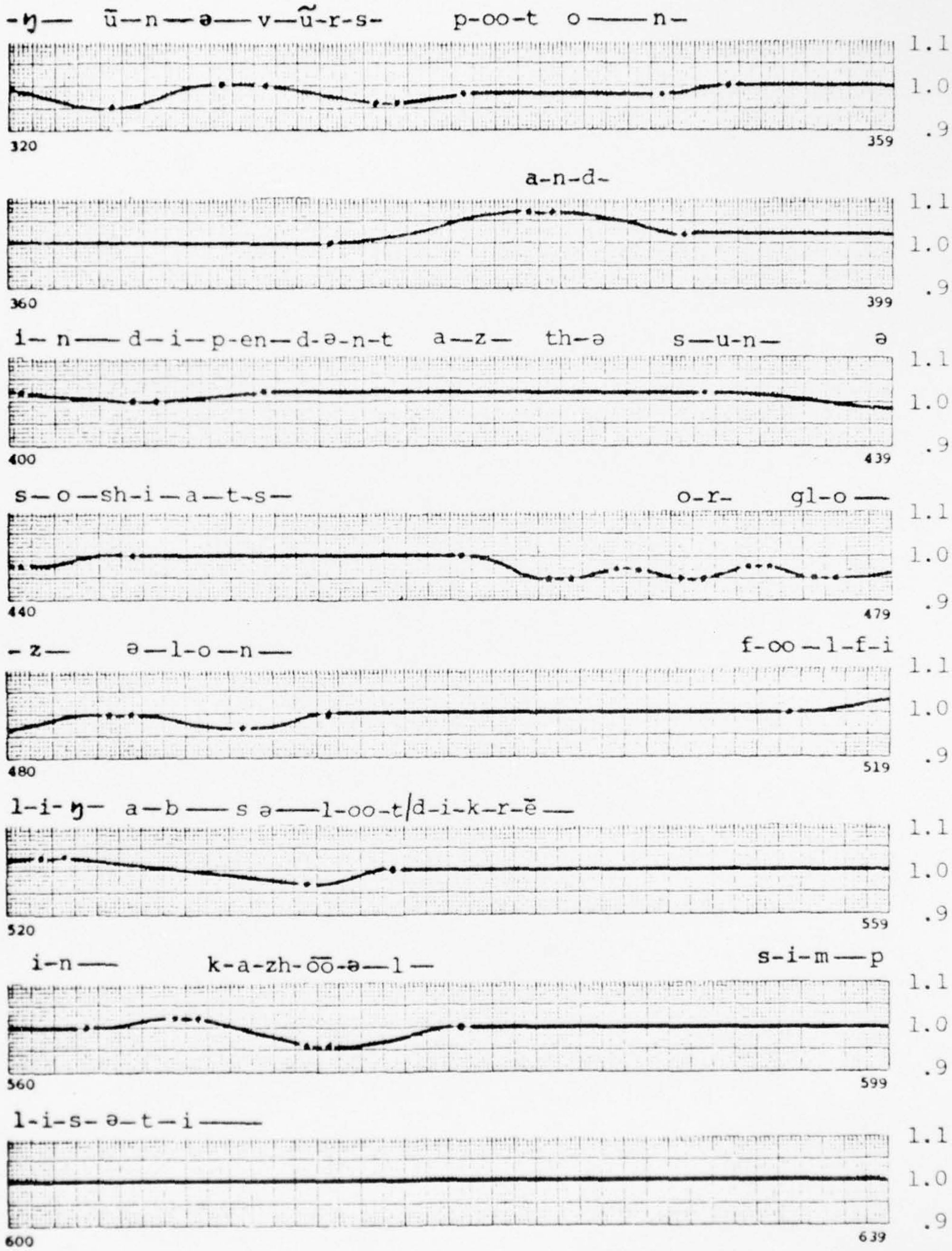


Figure B.12 (continued)

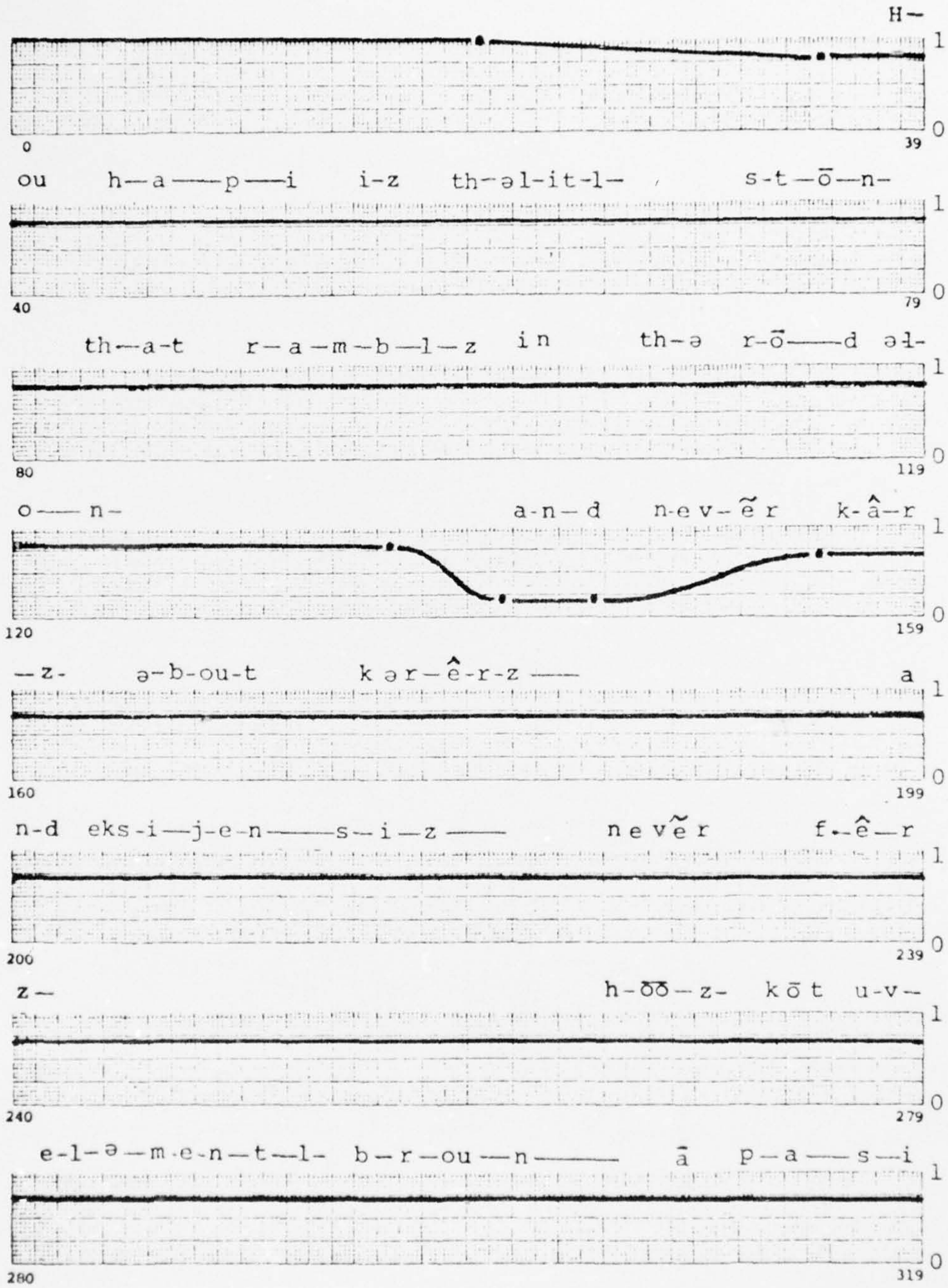


Figure B.13 - The expression parameter values used in sequence 1. A value of 1 is neutral and a value of 0 is a smile. The parameter value is .7 for frames 320 to 639.

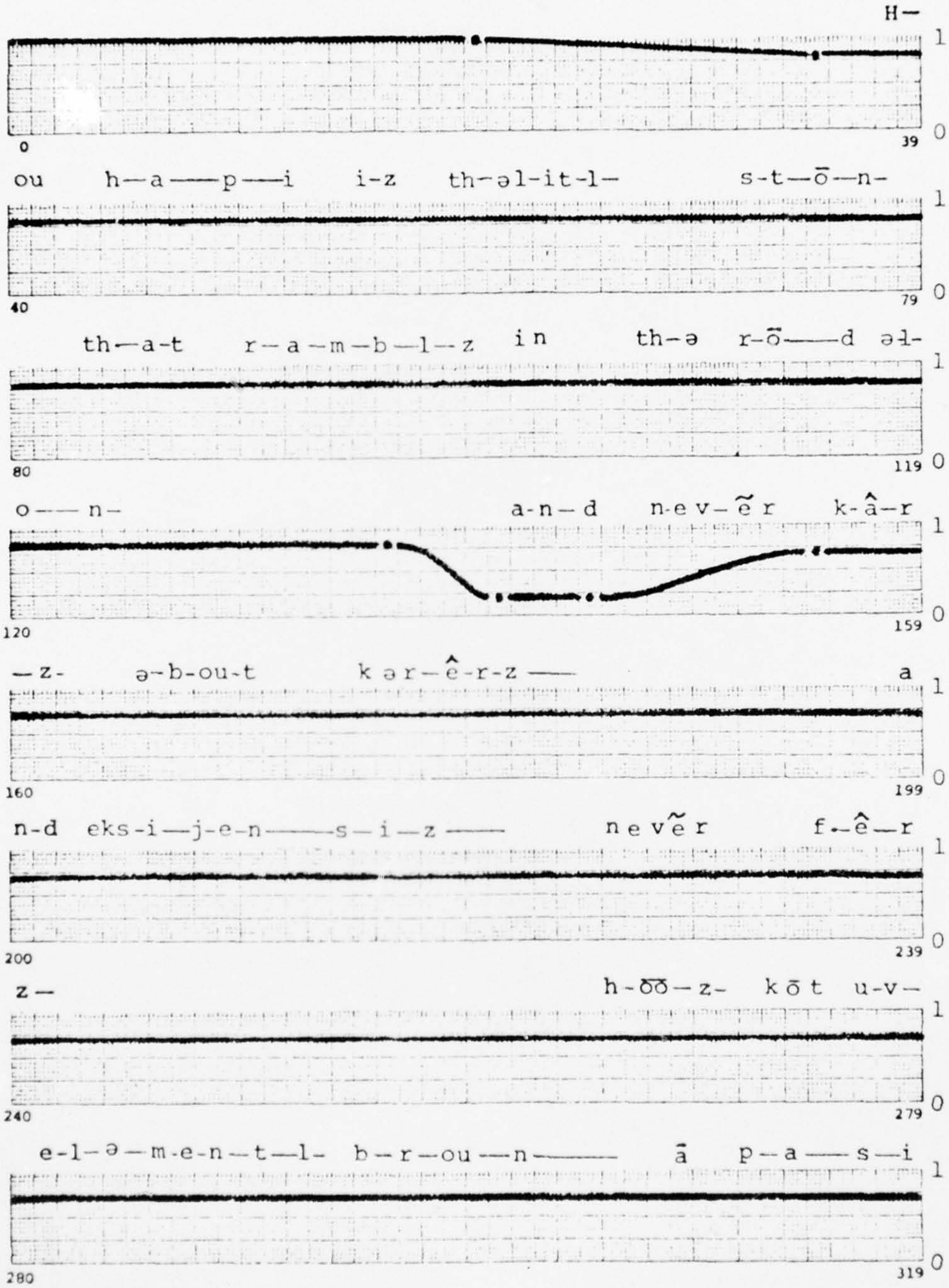


Figure B.14 - The expression parameter values used in sequences 2 through 5. For sequence 2 the parameter value is .7 in frames 320 to 639. A value of 1 is neutral and a value of 0 is a smile.

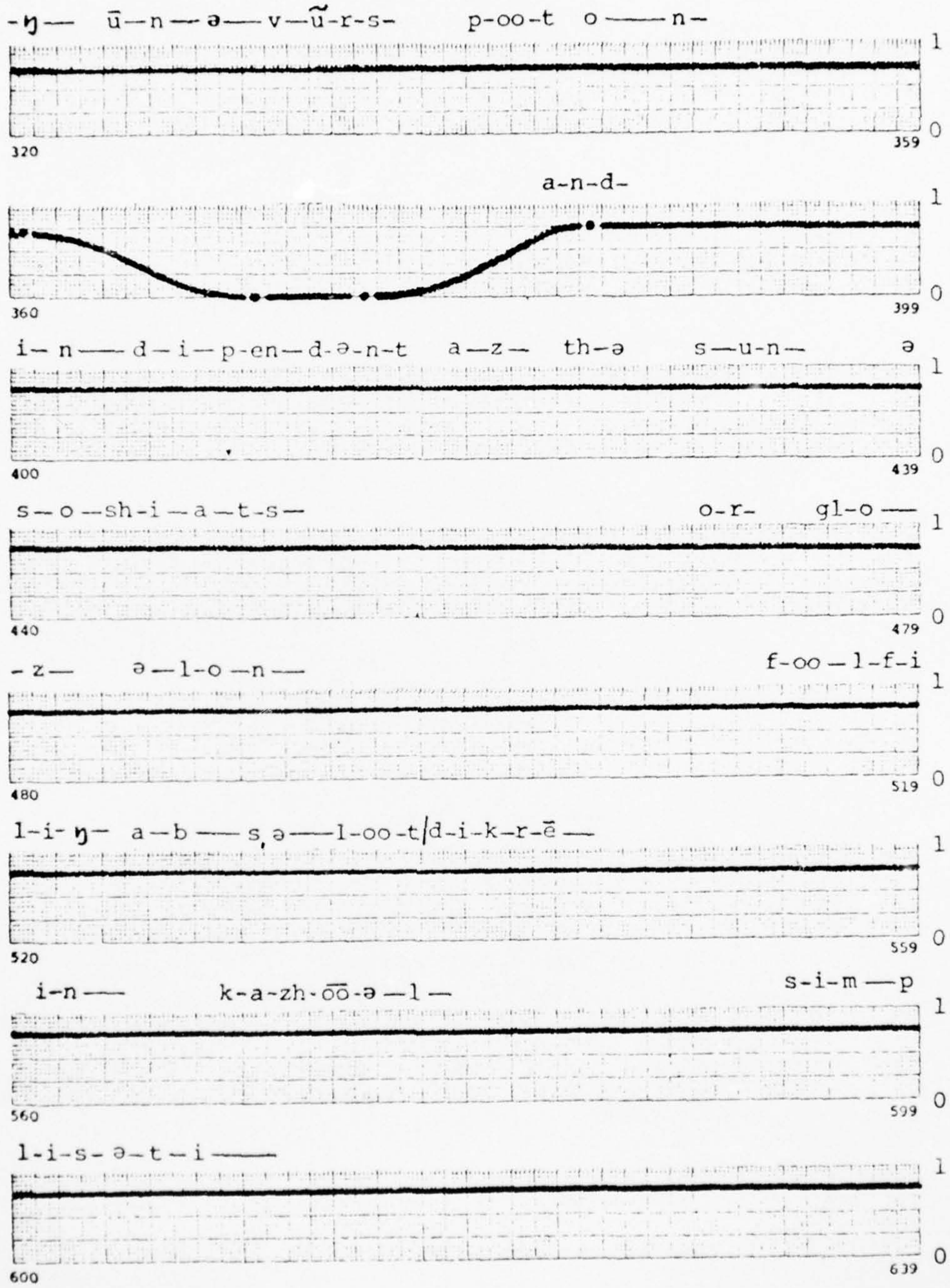


Figure B.14 (continued)

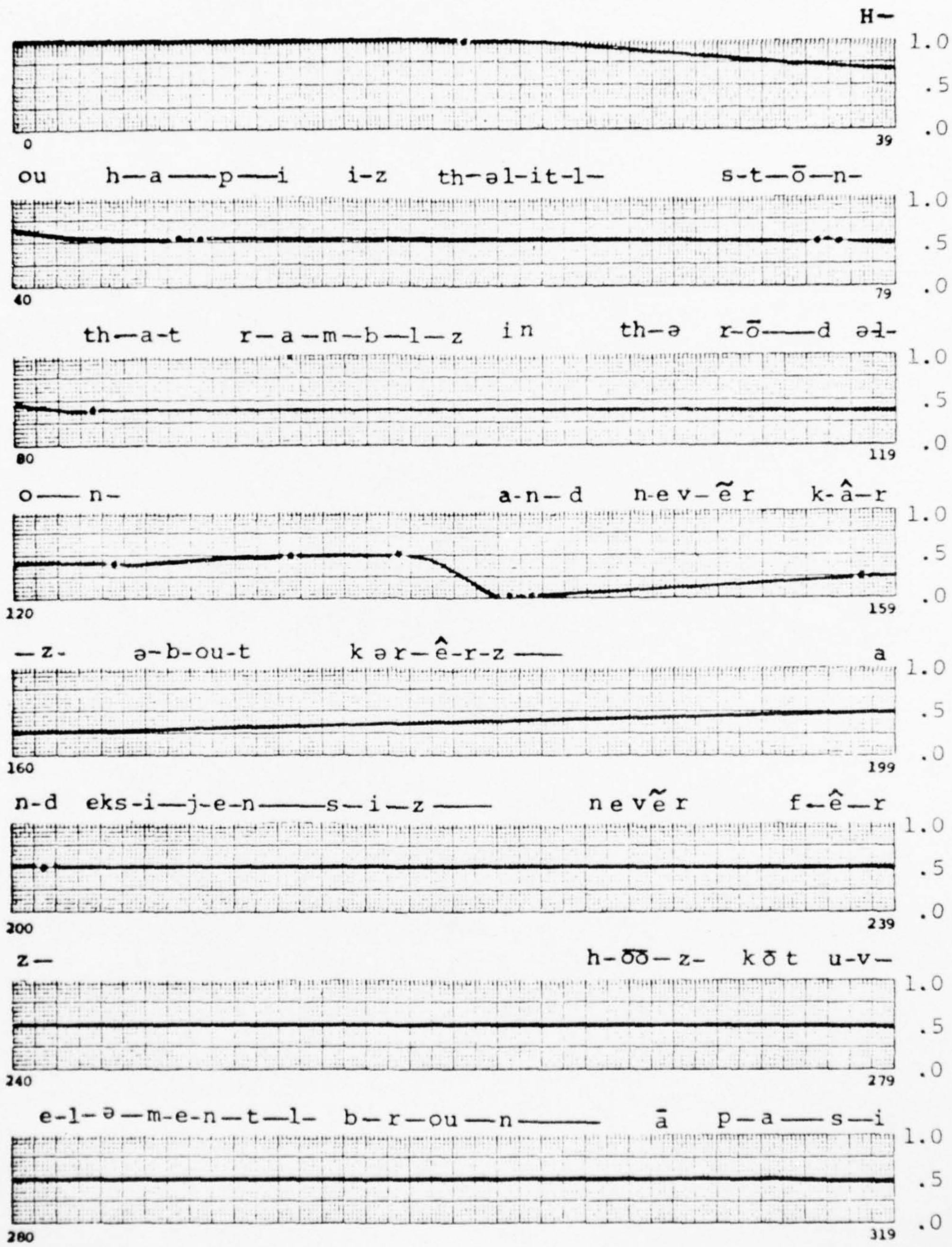


Figure B.15 - The expression parameter values used in sequence 6. A value of 1 is neutral and a value of 0 is a smile.

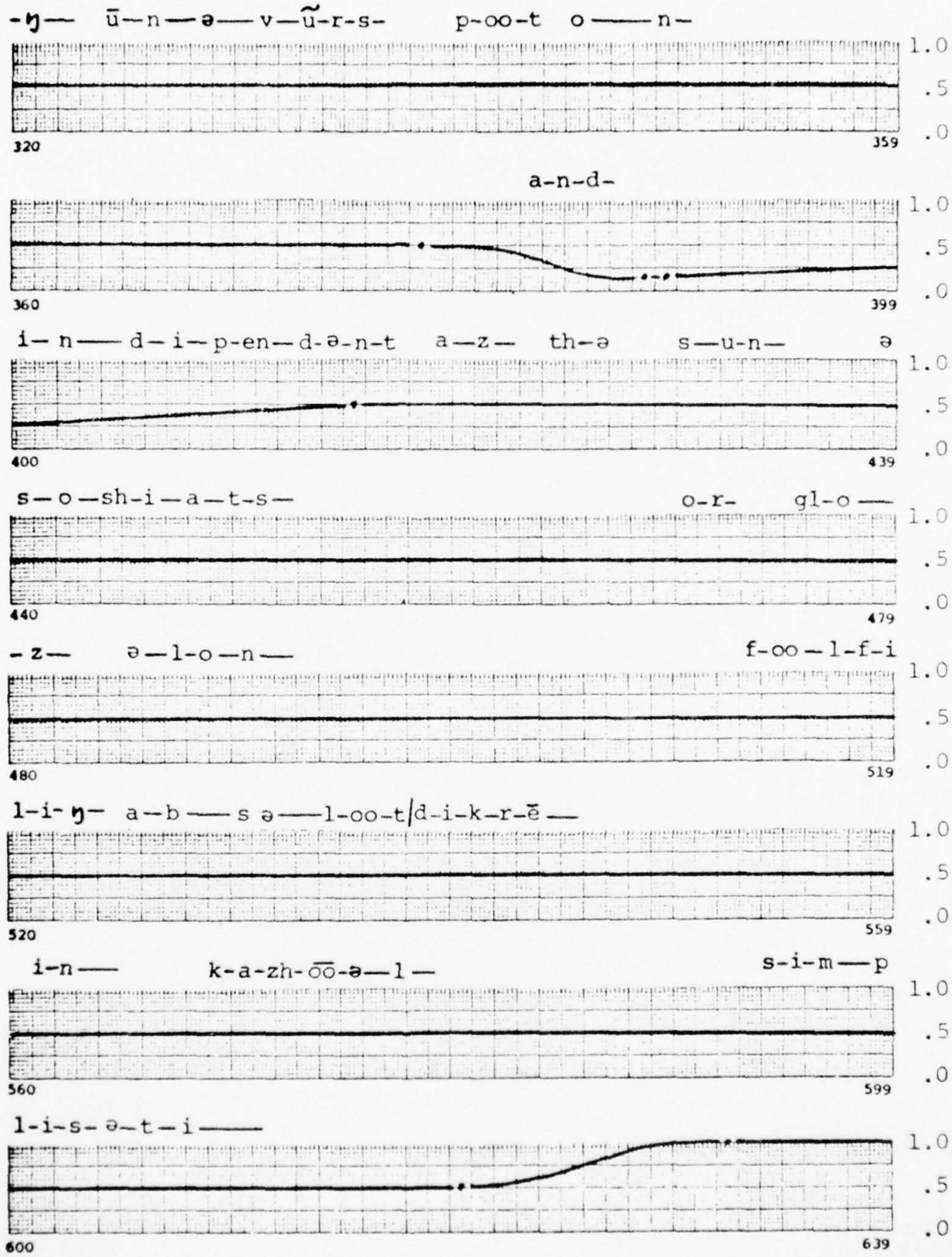


Figure B.15 (continued)

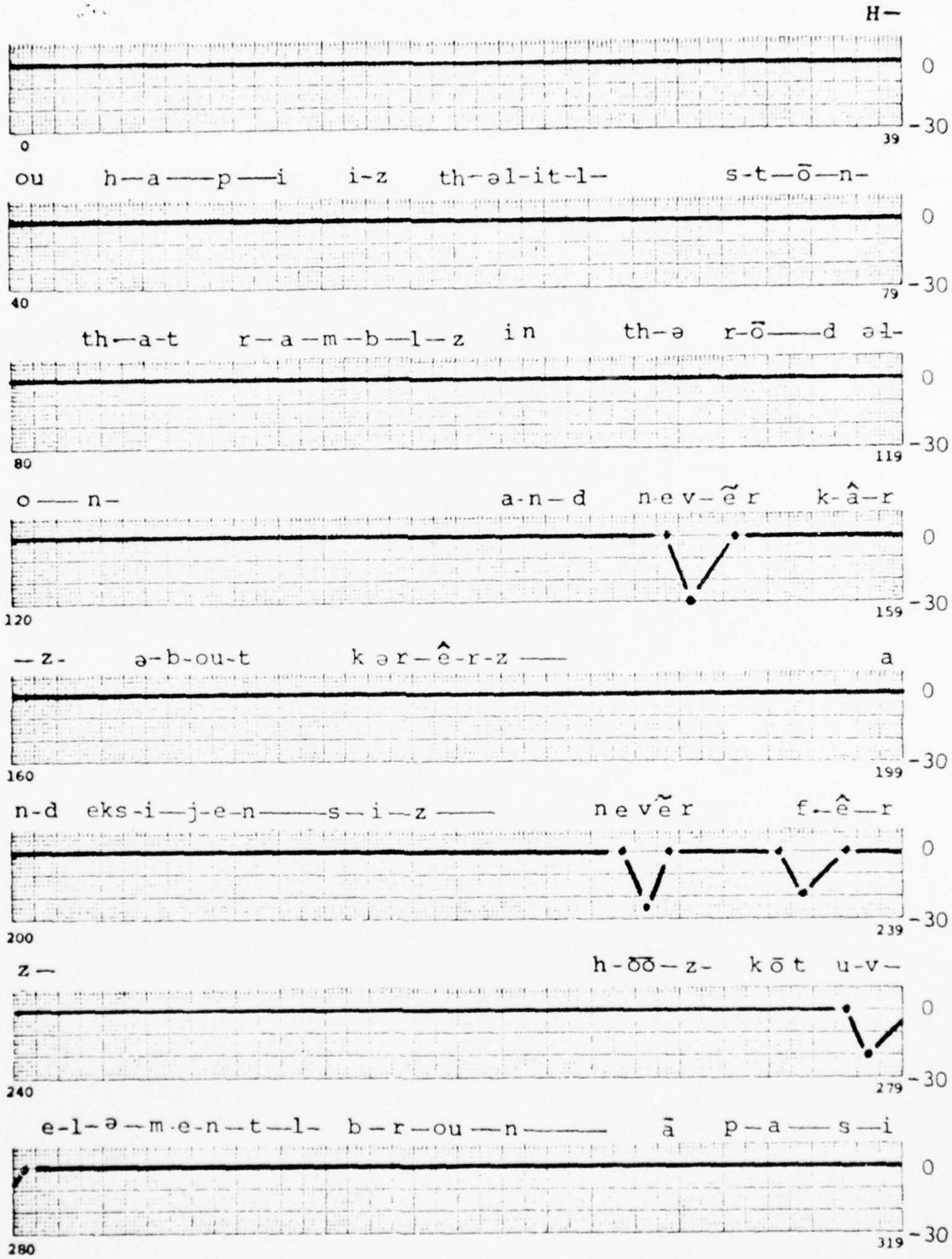


Figure B.16 - The f and v tuck parameter values used in sequence 1. A value of 0 corresponds to no tuck and a value of -30 corresponds to a full tuck.

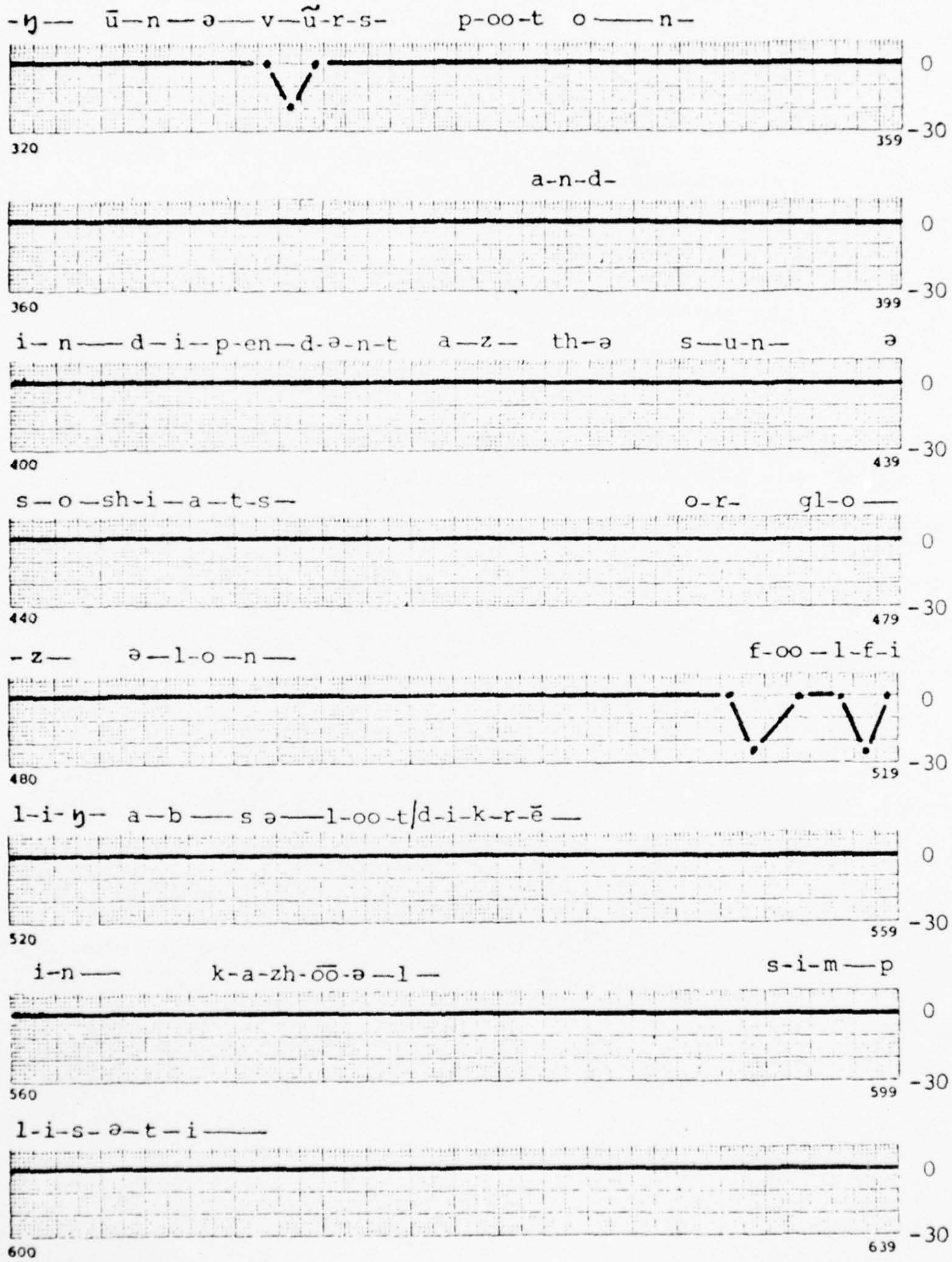


Figure B.16 (continued)

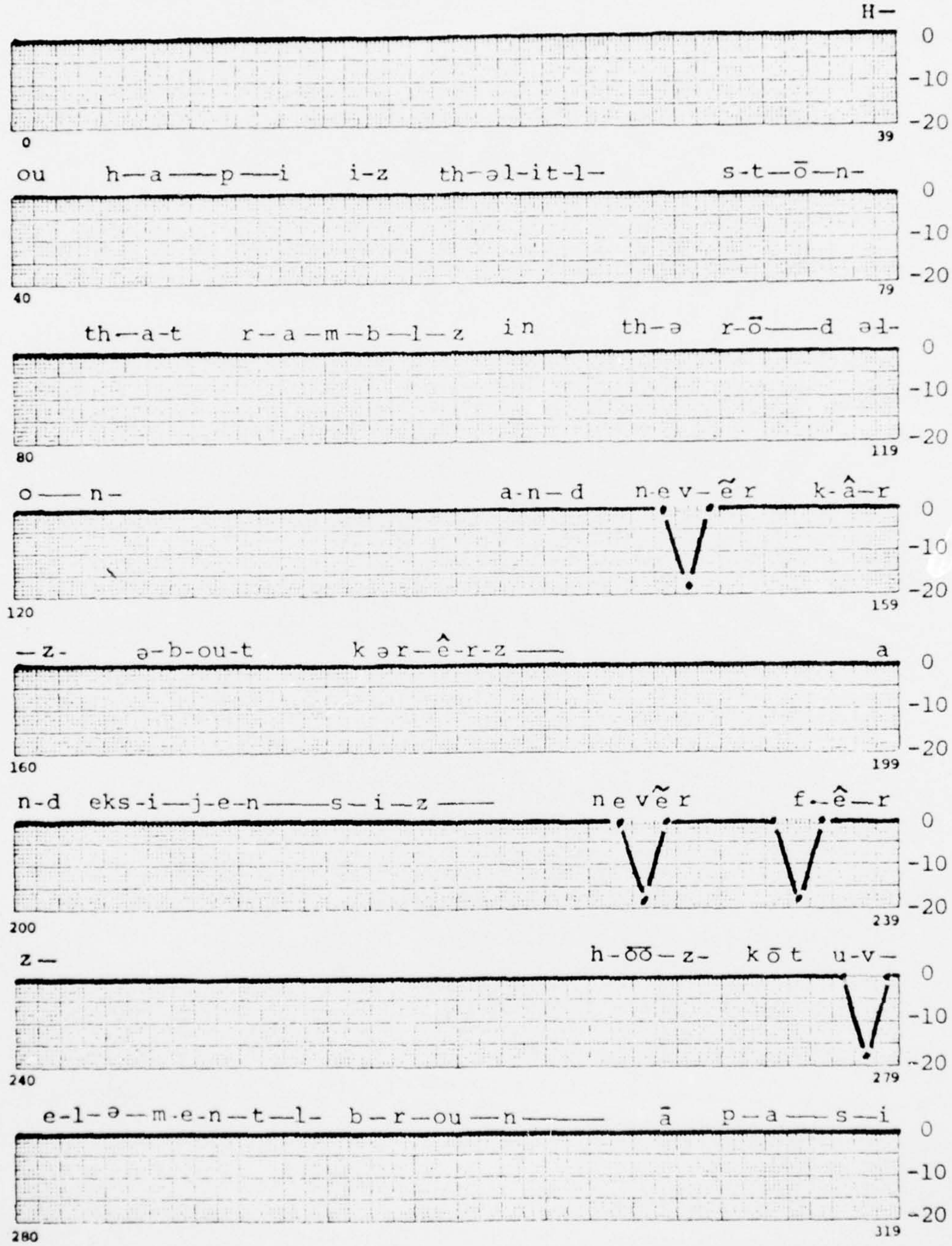


Figure B.17 - The f,v tuck parameter values used in sequences 3 and 4. A value of 0 corresponds to no tuck and a value of -20 corresponds to a full tuck.

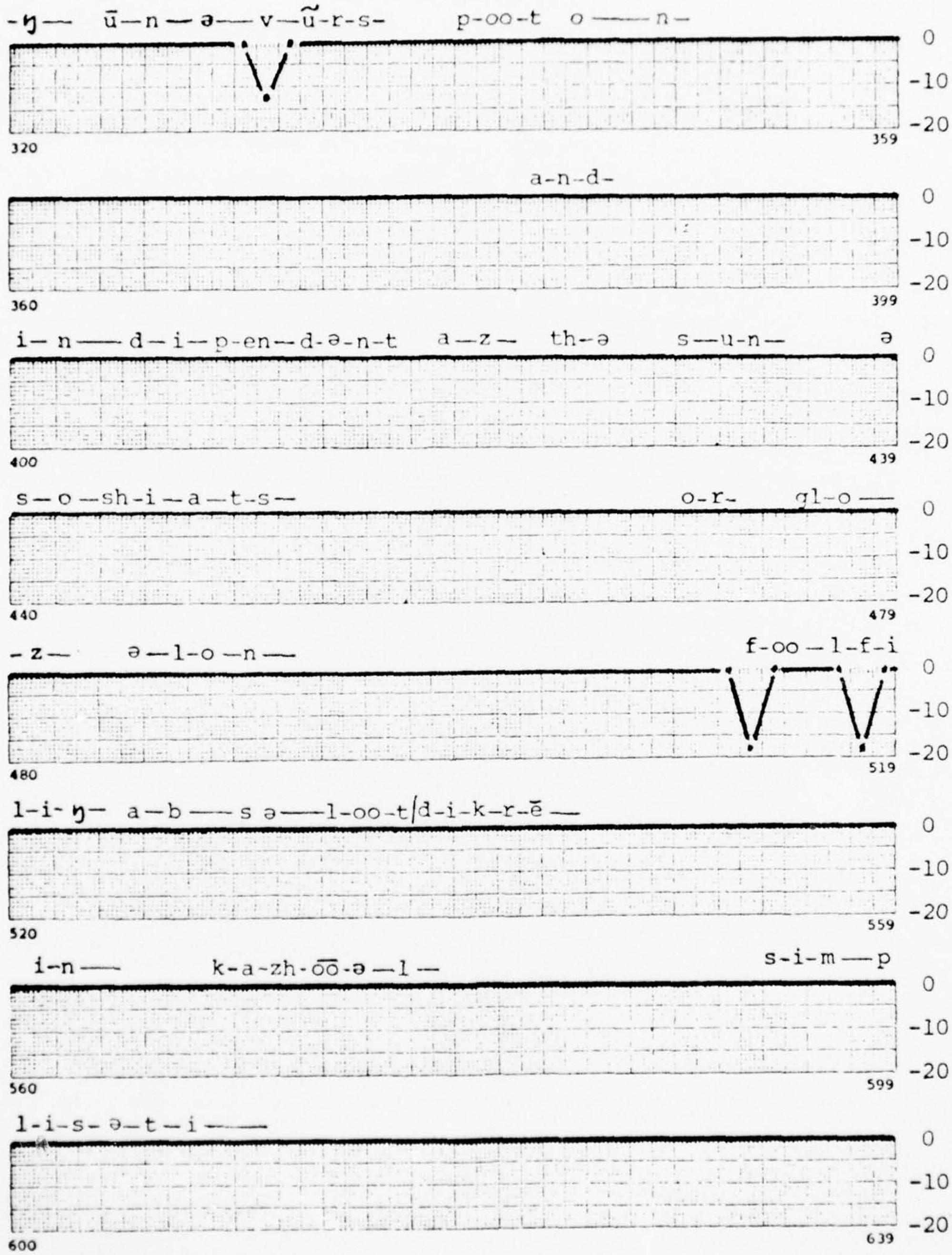


Figure B.17 (continued)

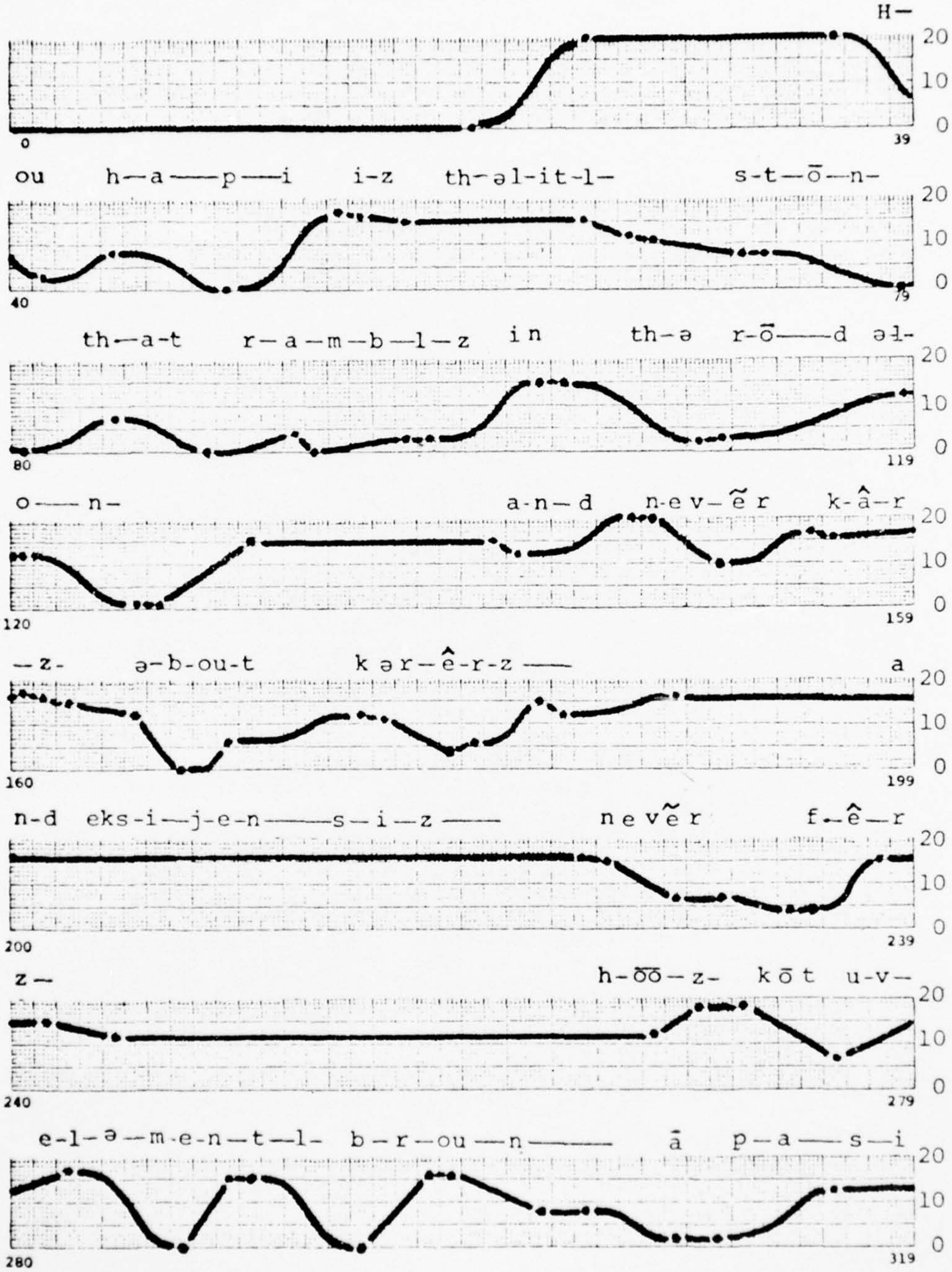


Figure B.18 - The upper lip parameter values used in sequence 5. A value of 20 corresponds to a fully raised lip while a value of 0 means the lip is not raised at all.

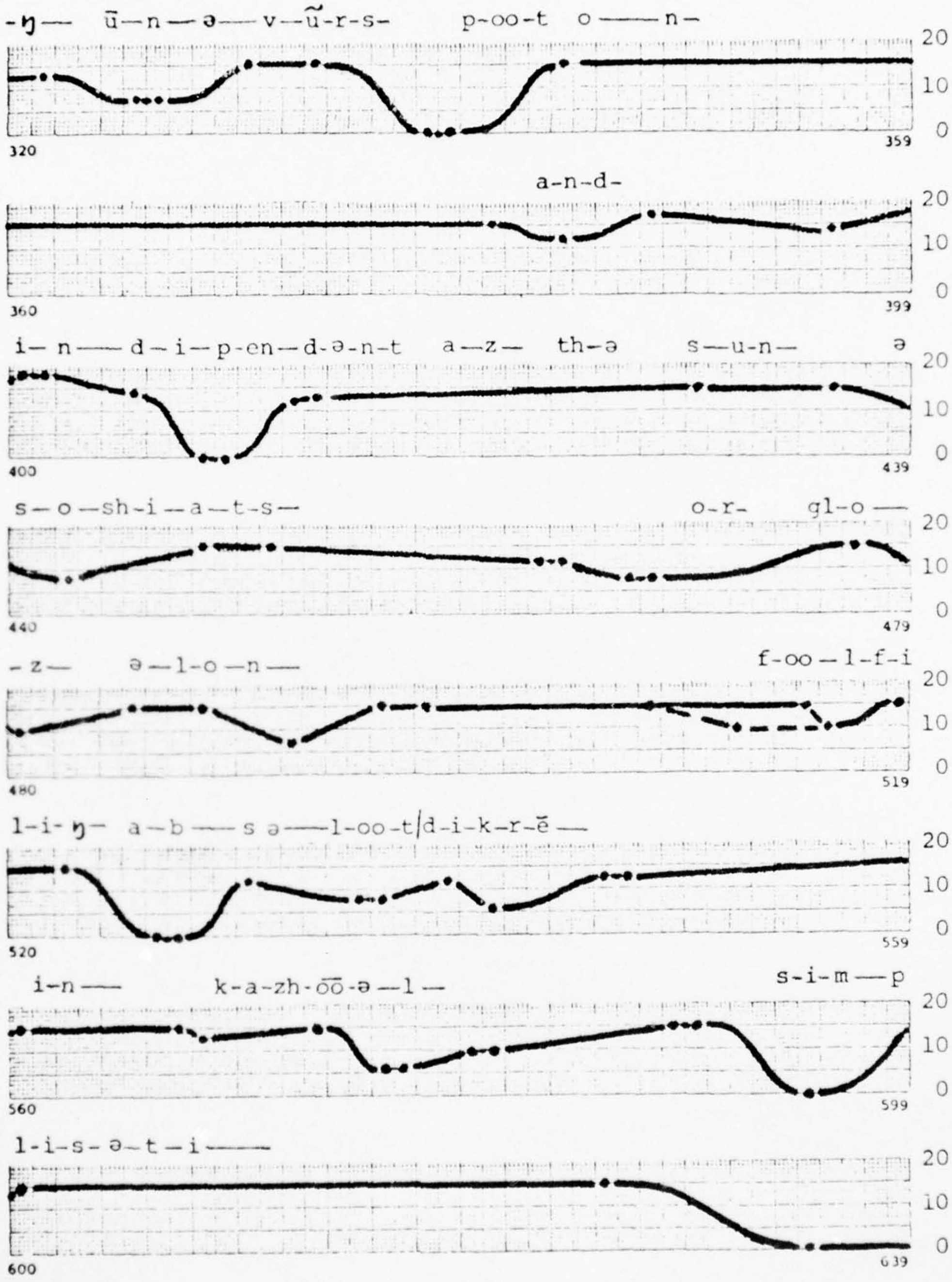


Figure B.18(continued) - In frames 508 to 516 the solid line is the actual function used. The dashed line indicates the intended function.

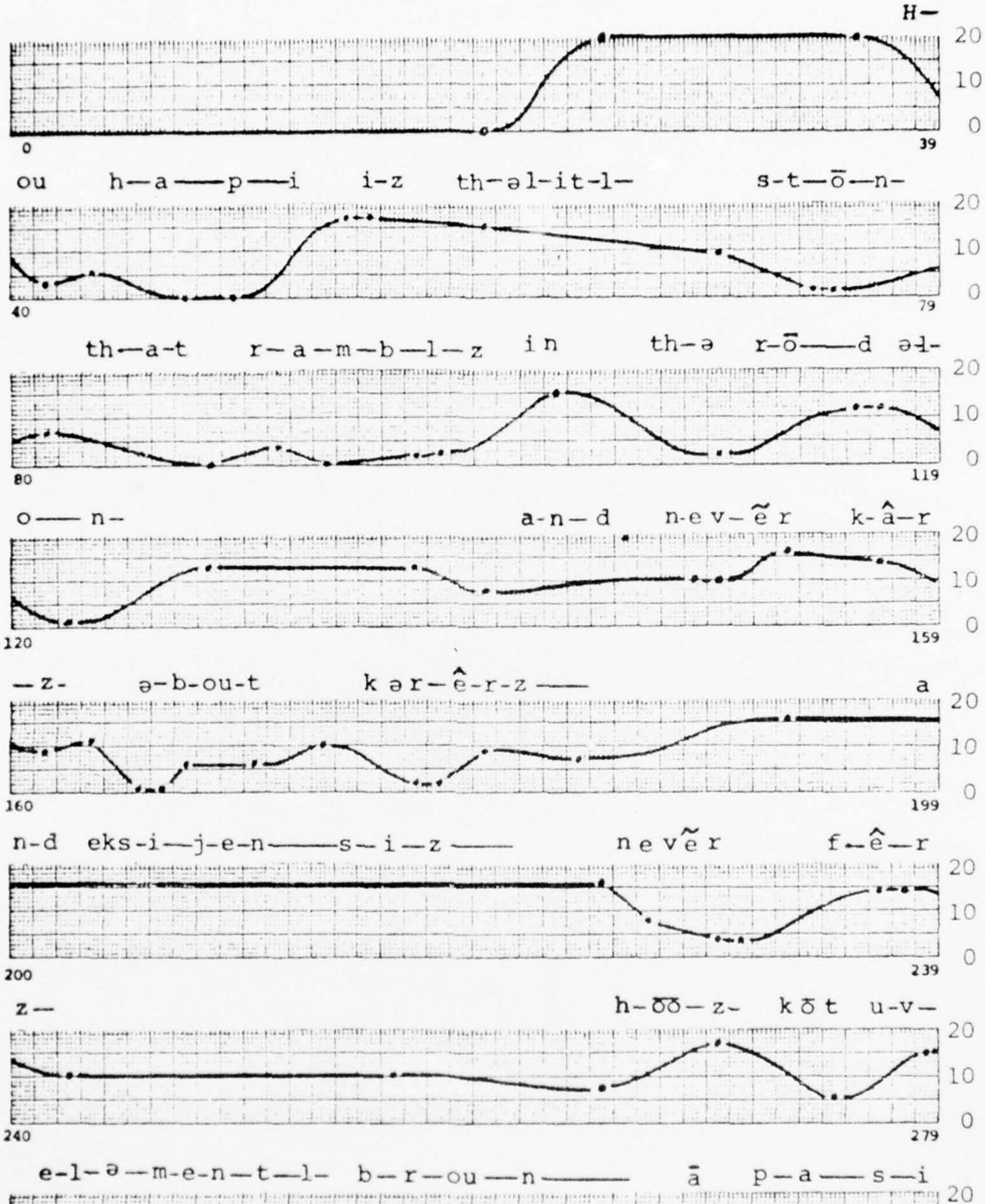


Figure B.19 - The upper lip parameter values used in sequence 6. A value of 20 corresponds to a fully raised lip whil a value of 0 means the lip is not raised at all.

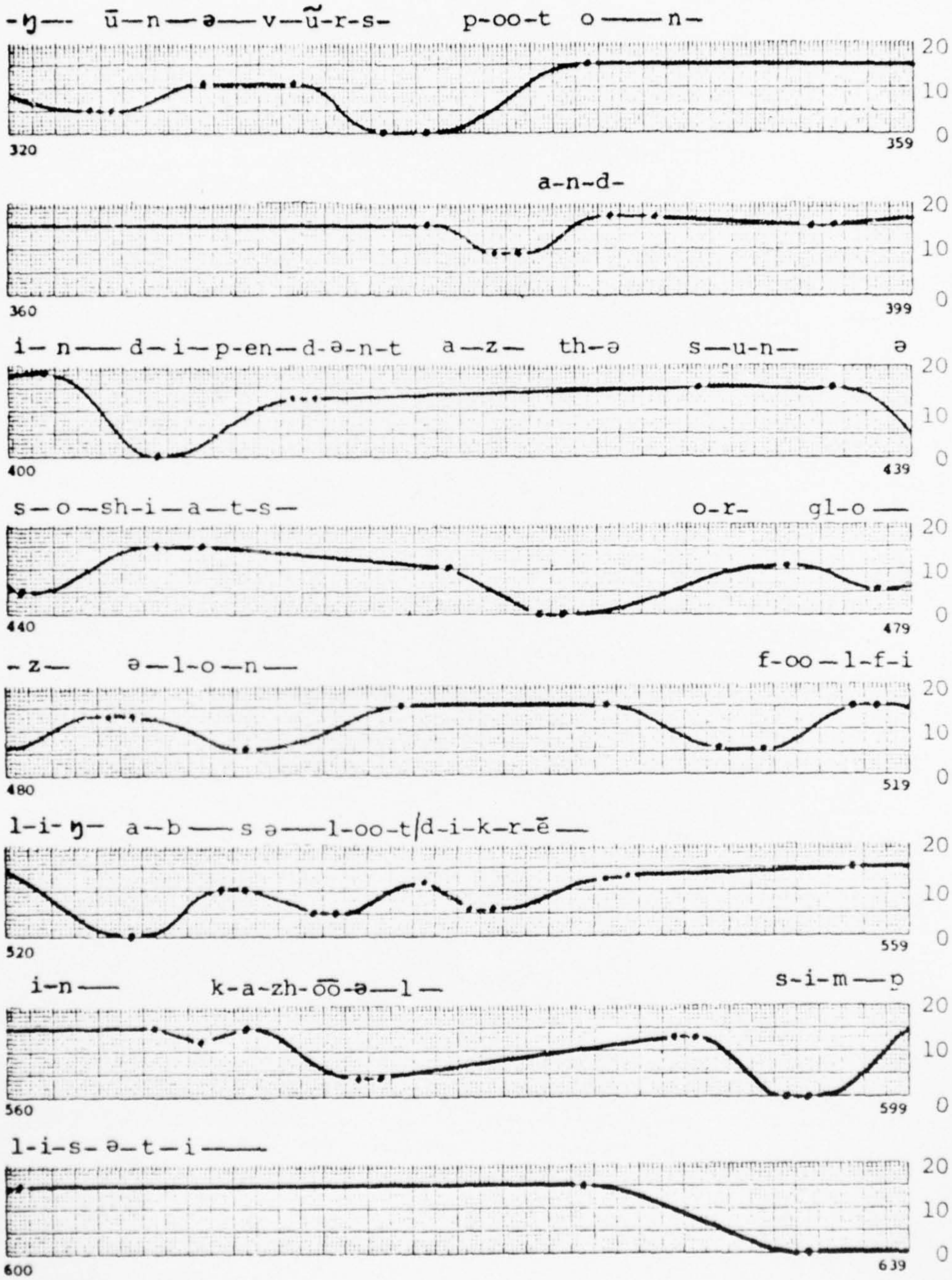


Figure B.19 (continued)

APPENDIX C

MEASURING THREE-DIMENSIONAL SURFACES WITH A TWO-DIMENSIONAL DATA TABLET

The measurement of point positions on a three-dimensional surface is a difficult task. This is particularly true if one does not have access to any of the specialized equipment used for this purpose and especially if the surface is flexible and varies over time. Outlined below is a technique that allows one to accomplish this task armed only with standard photographic techniques, a two-dimensional data tablet or digitizer and some computing power. The tablet or digitizer is not absolutely essential but is necessary for the method to be practical.

In essence the technique consists of photographing the surface of interest from several positions (simultaneously in the case of surfaces that vary over time) and then using the tablet to extract information from these multiple views of the surface. This information is then processed to determine the three-dimensional position of points on the surface.

C.1 Theory

A camera can be viewed as simply a device for projecting a three-dimensional space into a two-dimensional space. This mapping can be described mathematically using homogeneous coordinates

AD-A038 695 UTAH UNIV SALT LAKE CITY DEPT OF COMPUTER SCIENCE F/G 9/2

A PARAMETRIC MODEL FOR HUMAN FACES. (U)

DEC 74 F I PARKE DAHC15-73-C-0363

UNCLASSIFIED UTEC-CSC-75-047 NL

2 OF 2

AD A038695

END

DATE FILMED 5-77

techniques [12,13] and the proper mapping or transformation matrix.

$$[x\ y\ z\ 1]\begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{21} & T_{22} & T_{23} \\ T_{31} & T_{32} & T_{33} \\ T_{41} & T_{42} & T_{43} \end{bmatrix} = [x'\ y'\ w]$$

$$= w[u\ v\ 1] \qquad (1)$$

Where $u=x'/w$ and $v=y'/w$ are the coordinate values in the two-dimensional space of the photographs and x, y and z are the coordinate values in the three-dimensional space. Carrying out the indicated multiplication gives

$$T_{11}*x + T_{21}*y + T_{31}*z + T_{41} = x' = w*u$$

$$T_{12}*x + T_{22}*y + T_{32}*z + T_{42} = y' = w*v$$

$$T_{13}*x + T_{23}*y + T_{33}*z + T_{43} = w \qquad (2)$$

Substituting the expression for w into the first two equations above and collecting terms gives

$$(T_{11}-T_{13}*u)*x+(T_{21}-T_{23}*u)*y+(T_{31}-T_{33}*u)*z+(T_{41}-T_{43}*u)=0$$

$$(T_{12}-T_{13}*v)*x+(T_{22}-T_{23}*v)*y+(T_{32}-T_{33}*v)*z+(T_{42}-T_{43}*v)=0 \qquad (3)$$

If the three-space coordinates x, y, z and the transformation matrix T are known then these equations give the values of u and v, the two-dimensional coordinates in the photograph.

If the transformation matrix T and the two-space coordinates u and v are known then the equations above have the form

$$a*x + b*y + c*z + d = 0 \tag{4}$$

the equation of a plane in three-space. Now, if we have for the same point a u,v coordinate pair and the T matrix from another photographic view we will have four equations in the three unknowns x, y and z. In general if we have n views with n T matrices and n u,v coordinate pairs we will have 2n equations in the three unknowns.

$$\begin{vmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ & \cdot & \\ & \cdot & \\ & \cdot & \\ a_{2n} & b_{2n} & c_{2n} \end{vmatrix} \begin{vmatrix} x \\ y \\ z \end{vmatrix} = \begin{vmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_{2n} \end{vmatrix} \quad \text{or } [\ A\ ][\ X\ ] = [\ D\ ] \tag{5}$$

If a solution exists for this overdetermined system it can be computed in the following manner. Multiplying both sides of equation (5) by A^T gives

$$[\ A^TA\ ][\ X\ ] = [\ A^TD\ ] \tag{6}$$

This system of three equations when solved by standard techniques gives a least-mean-square solution for x, y and z. So, if we have the u,v coordinates of a point in at least two different photographic views and the corresponding transformation matrices for the views, it is possible to solve for the point's position in three-space.

Now, how does one determine the transformation matrix T for

each view? Going back to equation (1) we see that the transformation matrix contains 12 unknown values. But since we are dealing with a homogeneous system, the matrix will include an arbitrary scale factor and we are free to set one of the unknowns to any non-zero value. To find the 11 remaining unknowns we will need 11 equations of the form shown in (3) above. Since each point has two of these equations associated with it, the x,y,z and u,v coordinates for at least 5 1/2 points in each picture must be known in order to solve for T. If we use six points a,b,c,d,e,f and set T_{43}=1 the resulting system can be written in matrix form as follows.

$$\begin{vmatrix} x_a & y_a & z_a & 1 & 0 & 0 & 0 & 0 & -u_a*x_a & -u_a*y_a & -u_a*z_a \\ x_b & y_b & z_b & 1 & 0 & 0 & 0 & 0 & -u_b*x_b & -u_b*y_b & -u_b*z_b \\ x_c & y_c & z_c & 1 & 0 & 0 & 0 & 0 & -u_c*x_c & -u_c*y_c & -u_c*z_c \\ x_d & y_d & z_d & 1 & 0 & 0 & 0 & 0 & -u_d*x_d & -u_d*y_d & -u_d*z_d \\ x_e & y_e & z_e & 1 & 0 & 0 & 0 & 0 & -u_e*x_e & -u_e*y_e & -u_e*z_e \\ x_f & y_f & z_f & 1 & 0 & 0 & 0 & 0 & -u_f*x_f & -u_f*y_f & -u_f*z_f \\ 0 & 0 & 0 & 0 & x_a & y_a & z_a & 1 & -v_a*x_a & -v_a*y_a & -v_a*z_a \\ 0 & 0 & 0 & 0 & x_b & y_b & z_b & 1 & -v_b*x_b & -v_b*y_b & -v_b*z_b \\ 0 & 0 & 0 & 0 & x_c & y_c & z_c & 1 & -v_c*x_c & -v_c*y_c & -v_c*z_c \\ 0 & 0 & 0 & 0 & x_d & y_d & z_d & 1 & -v_d*x_d & -v_d*y_d & -v_d*z_d \\ 0 & 0 & 0 & 0 & x_e & y_e & z_e & 1 & -v_e*x_e & -v_e*y_e & -v_e*z_e \\ 0 & 0 & 0 & 0 & x_f & y_f & z_f & 1 & -v_f*x_f & -v_f*y_f & -v_f*z_f \end{vmatrix} \begin{vmatrix} T_{11} \\ T_{21} \\ T_{31} \\ T_{41} \\ T_{12} \\ T_{22} \\ T_{32} \\ T_{42} \\ T_{13} \\ T_{23} \\ T_{33} \end{vmatrix} = \begin{vmatrix} u_a \\ u_b \\ u_c \\ u_d \\ u_e \\ u_f \\ v_a \\ v_b \\ v_c \\ v_d \\ v_e \\ v_f \end{vmatrix}$$

or

$$[\ A\ ][\ T\ ] = [\ B\ ] \qquad (7)$$

If more than 5 1/2 points are available then multiplying both sides of (7) by A^T

$$[\, A^T A \,][\, T \,] = [\, A^T B \,] \tag{8}$$

and solving the resulting system of 11 equations using standard techniques will give a least-mean-square solution for the unknown T values. For those interested, an equivalent but more awkward approach to finding the T matrix is given in [1].

C.2 Application

You will have noted from the above discussion that this technique requires that each photograph must show at least six "reference" points whose three-dimensional x,y,z coordinates are known. This requirement has been satisfied by surrounding the surface to be measured with a cube of known dimensions. The corner points of the cube act as the refernce points. A wooden stick figure cube was build for small objects and a larger cube for human faces was constructed using weighted strings with beads forming the corners of the cube. See Figure C.1.

Since it is necessary to digitize exactly the same surface points in each of the photographs, some method of marking or identifying the surface points is needed. For irregular surfaces this may not be a problem but for smooth surfaces it was necessary to draw a grid of points on the surface as shown in Figure C.1.

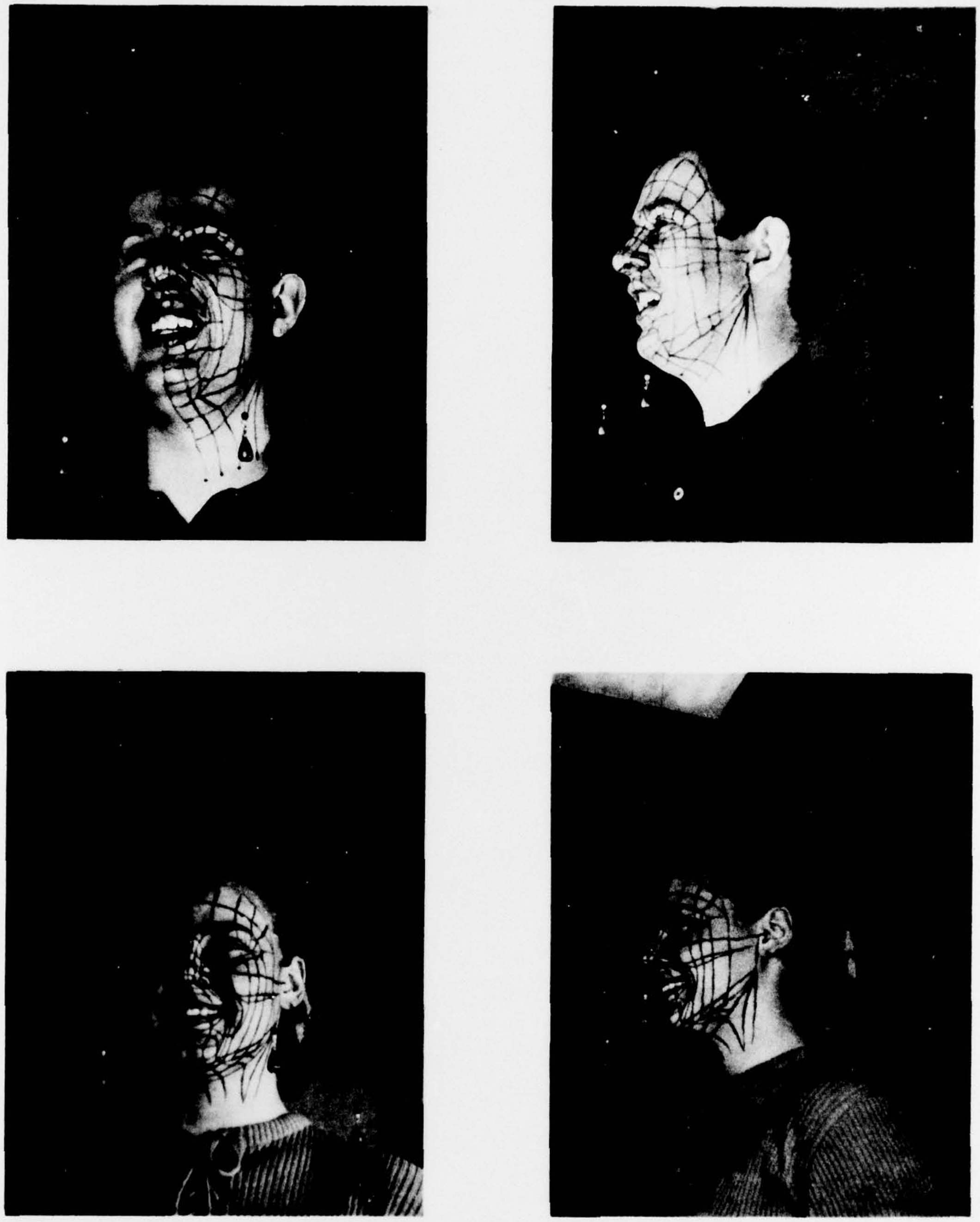


Figure C.1 - Data photographs.

C.3 Accuracy

Errors may enter the process in several ways. Obviously any error in determining the three-dimensional position of the reference points will influence the results. Errors in measuring the two-dimensional photograph coordinates of the points will also influence the results. In addition, computational errors (round off, truncation, etc.) may influence the results.

To minimize the errors in measuring the photograph coordinates, the photographs should be enlarged to take advantage of the maximum tablet resolution.

When solving the systems of equations (6) and (8) above, good numerical techniques such as Gaussian Elimination with partial pivoting should be used.

Some care should be taken in selecting camera positions for the various views. Every point of interest on the surface must be visible in at least two views. The camera positions should not be close to each other. As the camera positions approach each other small errors tend to be magnified.

A detailed error analysis has not been done but experience with the method indicates that with reasonable care in measuring the reference points and in digitizing the photographs the total error is less than one percent of the reference cube size. Typically the error is .5 percent or .05 inches when using a 10 inch reference cube. These figures apply when using 11 by 14 inch enlargements

Figure C.2 - Shaded images generated using data extracted from the data photographs shown in Figure C.1.

digitized on a tablet with .05 inch accuracy.

C.4 Conclusion

This technique has been successfully used to obtain point position data for the polygonal representation of several human faces and a human body. Figure C.2 shows several halftone images generated using point position data extracted from the photographs shown in Figure C.1.

APPENDIX D

THE SPEECH ANIMATION FILM

This appendix consists of a 16mm sound film titled "Speech Synchronized Computer Generated Facial Animation." The film contains the six computer generated speech animation segments discussed in Chapter 4. In addition it has a segment of the actor speaking. The parameter functions used to produce the computer generated segments are shown in Appendix B.

REFERENCES

1. Parke, F. I. Computer generated animation of faces. UTEC-CSc-72-120, Dept. of Computer Science, University of Utah, Salt Lake City (June 1972).

2. Parke, F. I. Computer generated animation of faces. Proc. ACM Annual Conference (August 1972), 451-457.

3. Watkins, G. S. A real-time visible surface algorithm. UTECH-CSc-70-101, Dept. of Computer Science, University of Utah, Salt Lake City (June 1970).

4. Gouraud, H. Computer display of curved surfaces. UTEC-CSc-71-113, Dept. of Computer Science, University of Utah, Salt Lake City (June 1971).

5. Gillenson, M. L. The interactive generation of facial images on a CRT using a heuristic strategy. Ph.D Dissertation, Ohio State Univ., Columbus, Ohio (1974).

6. Chernoff, H. The use of faces to represent points in n-dimensional space graphically. Technical Report 71, Project NR-042-993, Office of Naval Research (December 1971).

7. Sutherland, I. E. Three-dimensional data input by tablet. Proc. IEEE 62 (April 1974), 453-461.

8. Burton, R. P. Real-time measurement of multiple three-dimensional positions. UTEC-CSc-72-122, Dept. of Computer Science, University of Utah, Salt Lake City (June 1973), 51-60.

9. Private discussion with Ivan Sutherland and Ron Resch (September 1973).

10. Madsen, R. Animated Film: Concepts, Methods, Uses. Interland, New York, 1969, 120-122.

11. Webster's New World Dictionary of the American Language. World Publishing, New York, 1959.

12. Newman, W. M., and Sproull, R. F. Principles of Interactive Computer Graphics. McGraw-Hill, New York, 1973, 467-480.

13. Roberts, L. G. Homogeneous matrix representation and manipulation of n-dimensional constructs. MS 1405, MIT Lincoln Laboratory (May 1965).

ACKNOWLEDGMENTS

I am indebted to Ivan Sutherland for his continuing enthusiasm in this research and to Richard Riesenfeld for his careful reading of the dissertation draft.

I would like to thank Barr Cannon for helping me put together the speech animation film, Mike Milochik for his assistance in preparing the photographs for the dissertation, Ali Atashroo for helping me collect the face data, and Ron Resch for his helpful suggestions.

I addition, I would like to express my appreciation to Barry Wessler, Bui Tuong-Phong, Nelson Dinerstein, Debra Youngberg, James Clark, Robert Burton, Mike Cannon, Ivan Sutherland, Nina Saffari, Robert Ingebretsen and Ali Atashroo for "volunteering" their faces as surfaces to be measure.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER UTEC-CSc-75-047	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) A PARAMETRIC MODEL FOR HUMAN FACES		5. TYPE OF REPORT & PERIOD COVERED Technical Report
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Frederick Ira Parke		8. CONTRACT OR GRANT NUMBER(s) DAHC15-73-C-0363
9. PERFORMING ORGANIZATION NAME AND ADDRESS Computer Science Department University of Utah Salt Lake City, Utah 84112		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order #2477
11. CONTROLLING OFFICE NAME AND ADDRESS Defense Advanced Research Projects Agency 1400 Wilson Blvd. Arlington, Virginia 22209		12. REPORT DATE December 1974
		13. NUMBER OF PAGES 106
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

This document has been approved for public release and sale; its distribution is unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

3-dimensional shaded graphics, parametric model, computer animation, measurement of 3-dimensional surfaces, polygonal model, synchronized animation, representation of human faces

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report presents a computer model for the representation of human faces. This three-dimensional, parametric model produces shaded facial images. The face, constructed of polygonal surfaces, is manipulated through the use of parameters which control interpolation, translation, rotation and scaling of the various facial features.

With this model, very little input information is needed to specify and generate a specific face with a specific expression. The model has been

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

successfully used to produce a large variety of facial images and several animated sequences. The animated sequences illustrate the power of the model to change facial expression and conformation.

Experience with the model indicates that fewer than 10 parameters must be manipulated to produce reasonable speech synchronized facial animation.